https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022


Biogeosciences Discussions
Open Access
EGU

# Assimilation of multiple different datasets results in large differences in regional to global-scale NEE and GPP budgets simulated by a terrestrial biosphere model

Cédric Bacour[1 ◎], Natasha MacBean[2], Frédéric Chevallier[1], Sébastien Léonard[1 ※], Ernest. N. Koffi[1 ☆], Philippe Peylin[1]

[1] Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, F-91191, France.

[2] Department of Geography, Indiana University, Bloomington, IN, 47405, USA.

◎ formerly at NOVELTIS, Labège, France.

※ now at Air Liquide R&D, Innovation Campus Paris - Les-Loges-en-Josas, France.

☆ now at European Centre for Medium-Range Weather Forecasts, Robert-Schuman-Platz 3, 53175 Bonn, Germany.

*Correspondance to*: Cédric Bacour (cedric.bacour@lsce.ipsl.fr)

## Key Points:

- The impact of assimilating different dataset combinations on regional to global scale C budgets is explored with the ORCHIDEE model
- Assimilating simultaneously multiple datasets is preferable to optimize the values of the model parameters and avoid model overfitting
- The challenges in optimizing soil C pools using atmospheric $CO_2$ data are highlighted for an accurate prediction of the land sink distribution

## Abstract

In spite of the importance of land ecosystems in offsetting carbon dioxide emissions released by anthropogenic activities into the atmosphere, the spatio-temporal dynamics of the carbon fluxes remain largely uncertain at regional to global scales. Over the past decade, Data Assimilation (DA) techniques have grown in importance for improving these fluxes simulated by Terrestrial Biosphere Models (TBMs), by optimizing model parameter values while also pinpointing possible parameterization deficiencies. Although the joint assimilation of multiple data streams is expected to

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022

Biogeosciences Discussions
Open Access EGU

constrain a wider range of model processes, their actual benefits in terms of reduction in model uncertainty are still under-researched, also given the technical challenges. In this study, we investigated with a consistent DA framework and the ORCHIDEE-LMDz TBM-atmosphere model how the assimilation of different combinations of data streams may result in different regional to global carbon budgets. To do so, we performed comprehensive DA experiments where three datasets (*in situ* measurements of net carbon exchange and latent heat fluxes, space-borne estimates of the Normalized Difference Vegetation Index, and atmospheric $CO_2$ concentration data at stations) are assimilated alone or simultaneously. We thus evaluated their complementarity and usefulness to constrain net and gross C fluxes. We found that a major challenge in improving the spatial distribution of the land sinks/sources with atmospheric $CO_2$ data relates to the correction of the initial carbon stocks.

## 1 Introduction

The dramatic increase of atmospheric $CO_2$ concentrations recorded in the last half-century has grown awareness on the determining impact of human activities on climate. Taking up about one third of the carbon dioxide from the atmosphere, the terrestrial biosphere plays a key role in regulating $CO_2$ emissions released by anthropogenic activities (fossil fuel emissions, land use and land cover change) (Friedlingstein et al., 2020). Quantifying variations in the distribution and intensity of carbon sources/sinks from year to year remains a challenge given the complexity of the processes involved and what we can learn from observations. By formalizing current knowledge of the main processes governing the functioning of vegetation into numerical representations, terrestrial biosphere models (TBMs) have grown in importance for studying the spatio-temporal dynamics of net and gross land surface carbon (C) fluxes from the local to the global scales. However, the large spread in their simulated regional to global scale C fluxes for the present time (Friedlingstein et al., 2020) as well as for future projections (Arora et al., 2020) highlight the remaining uncertainties in our understanding and prediction of the fate and role of the biosphere under climate change and anthropogenic pressure.

Over the past decade, the parameter uncertainty in TBMs has increasingly been reduced thanks to statistical data assimilation (DA, also referred to as model-data fusion) frameworks, benefiting from the experience gained in other fields of Earth and Environmental sciences (geophysics, weather forecasting, hydrology, oceanography, etc.). DA technique enables optimizing the model parameters using relevant target observations, while taking into account both observational and modelling

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022




uncertainties. DA does not only enable improving the model parameters but can also help pinpointing model deficiencies (Luo et al., 2012). The importance of DA as a key component of terrestrial biosphere carbon cycle modelling is reflected by the diversity of DA systems in the global TBM communities. Since the first global scale Carbon Cycle Data Assimilation System (CCDAS) (Kaminski et al., 2002; Rayner et al., 2005) developed for the Biosphere Energy-Transfer Hydrology (BETHY) model, other modelling groups have developed global scale carbon cycle DA systems, in particular for ORCHIDEE (ORganizing Carbon and Hydrology In Dynamic EcosystEms model) (Santaren et al., 2007; Peylin et al., 2016), JULES (Joint UK Land Environment Simulator) (Raoult et al. (2016)), JSBACH (Schürmann et al. (2016)), or CLM (Community Land Model) (Fox et al., 2018).

Within a DA framework, ground-based measurements of eddy-covariance fluxes at a local scale (Wang et al., 2001; Knorr and Kattge, 2005; Sacks et al., 2007; Williams et al., 2009; Groenendijk et al., 2011; Kuppel et al., 2012) have been widely used to constrain net and gross $CO_2$ fluxes and latent heat flux. Moreover, remote sensing proxies of vegetation activities, such as raw reflectance data (Quaife et al., 2008), vegetation indices (Migliavacca et al., 2009; MacBean et al., 2015), or FAPAR - fraction of absorbed photosynthetically active radiation (Stöckli et al., 2008; Zobitz et al., 2014; Forkel et al., 2014; Bacour et al., 2015), have also been used to constrain the model parameters at various spatial scales. Finally, atmospheric $CO_2$ mole fraction measurements have been assimilated to provide valuable information on large-scale net ecosystem exchange (NEE) (Rayner et al., 2005; Koffi et al., 2012).

In the early days of DA studies, most focused on the assimilation of a single data stream (e.g., only NEE). Then, assimilations with multiple different C cycle related datasets have soon been considered (Moore et al., 2008; Richardson et al., 2010; Ricciuto et al., 2011; Keenan et al., 2013; Thum et al., 2017; Knorr et al., 2010; Kaminski et al., 2012; Kato et al., 2013; Bacour et al., 2015; Peylin et al., 2016). The underlying motivation behind assimilating multiple data streams is that using a greater number and diversity of observations should provide stronger constraints on model parameters, including a wider range of processes, hence resulting in a greater reduction in model uncertainty. However, many previous studies that assimilated multiple datasets hardly considered potential incompatibilities between the model and the observations (Bacour et al., 2015; Thum et al., 2017). Besides, a few have quantified the actual benefit of assimilating multiple data-sets compared to the single data stream assimilations, in particular in the context of global scale C cycle DA experiments.
The assimilation of multiple data streams can be done either sequentially, in which one observation type is assimilated at a time, or simultaneously (joint assimilation approach or "batch" strategy as defined in Raupach et al., 2005), where the model is calibrated with all data included in the same

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022




optimization (e.g. Richardson et al., 2010; Kaminski et al., 2013; Schürmann et al., 2016). While the latter joint assimilations are more optimal as it maximizes the consistency of the model with the whole of the datasets considered (Richardson et al., 2010), sequential approaches remain appealing for modelers: They require less initial technical investment and enable easier assessment of the impact of the data stream assimilated successively onto the optimized variables. Both approaches however face similar challenges, like defining the model-data uncertainty and hence the weight of that each dataset has on the optimization outcome (see, e.g., Richardson et al., 2010; Keenan et al., 2013; Kaminski et al., 2012; Bacour et al., 2015; Thum et al., 2017; Peylin et al., 2016). Another major challenge, as highlighted by MacBean et al. (2016), concerns inconsistencies between observations and model outputs, which are usually not accounted for in common bias-blind (Dee, 2005) Bayesian DA systems relying on the hypothesis of Gaussian errors. Indeed, most studies do not attempt to identify systematic errors in the observations and/or in the model and to correct for them. The likely impact of model-data biases on the parameter optimization is then a degraded model performance as well as an illusory decrease in the estimated model uncertainty (Wutzler and Carvalhais, 2014; MacBean et al., 2016; Bacour et al., 2019).

The present study aims to go a step forward in the assessment of how assimilating multiple different C cycle related data streams impacts the constraint on net and gross $CO_2$ flux simulations at the global scale. To do so, we further advance from the sequential assimilation of Peylin et al. (2016) (referred to as "stepwise" approach hereafter) by implementing a simultaneous assimilation framework with the same data streams: net carbon fluxes (net ecosystem exchange – NEE) and latent heat fluxes (LE) measured at eddy covariance sites across different ecosystems, satellite derived Normalized Difference Vegetation Index (NDVI) at coarse resolution for a set of pixels spanning the main deciduous vegetation types, and monthly atmospheric $CO_2$ concentration data measured at surface stations worldwide. The study relies on the DA framework designed for the ORCHIDEE global vegetation model (Krinner et al., 2005), here associated to a simplified version of the LMDz atmospheric transport model (Hourdin et al., 2006) based on pre-calculated transport fields for assimilating atmospheric $CO_2$ concentration data. ORCHIDEE and LMDz are the terrestrial and atmospheric components of the IPSL Earth System Model (Dufresne et al., 2013). As a preliminary step, we tuned prior and observation error statistics: we enhanced their realism by making them consistent with the differences between prior model simulations and observations. We then conducted different assimilation experiments in which each data stream was assimilated alone on in combination (for all combinations of datasets) to assess what the impact of each assimilation experiment was on: a) the fit to each dataset (including datasets included and excluded from the assimilations); b) on the resultant regional to global scale gross and net $CO_2$ flux budget (NEE and

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022



GPP); and c) on the optimized parameter values and uncertainties. We complemented our analysis by comparing our regional to global scale C budgets with independent process-based model and atmospheric inversion estimates from the Global Carbon Project's 2020 Global Carbon Budget (Friedlingstein et al., 2020).

We completed our assessment of the impact of each assimilation scenario by quantifying the observation influence of each of the three data streams on the joint assimilation in which all three datasets were included in the optimization. By determining the overall constraint brought by each data set on the inversion, this final analysis allowed us to highlight the importance of atmospheric $CO_2$ data in the optimisation of soil C pools despite the smaller number of observations assimilated. Throughout the presentation of the results, we discuss implications of each assimilation experiment on our ability to accurately constrain gross and net $CO_2$ fluxes. In the final section we propose some perspectives for other modeling groups wishing to implement global scale parameter DA systems to constrain regional to global scale C budgets.

# 2 Materials and methods

## 2.1 Models

### *2.1.1 ORCHIDEE*

Model description

ORCHIDEE is a spatially explicit process-based global TBM (Krinner et al. 2005) that calculates the fluxes of carbon dioxide, water and heat, between the biosphere and the atmosphere, as well as the soil water budget. The temporal resolution is half an hour except for the slow components of the terrestrial carbon cycle (including carbon allocation in plant reservoirs, soil carbon dynamics, and litter decomposition) which are calculated on a daily basis. The version of ORCHIDEE in this study corresponds to that used in the IPSL model for its contribution to the Climate Model Intercomparison Project 5 (CMIP5) established by the World Climate Research Program (https://cmip.llnl.gov/). Vegetation is represented by 13 Plant Functional Types (PFTs) that include bare soil. The processes use the same governing equations for all PFTs, except for the seasonal leaf dynamics (phenology), which follows Botta et al. (2000) (see MacBean et al. (2015) for a full description). The observation operator for NDVI is determined *i)* by assuming a linear relationship between NDVI and FAPAR (as in MacBean et al., 2015) and *ii)* by calculating FAPAR from the simulated LAI based on the classical Beer-Lambert law for the extinction of the direct illumination within the canopy (Bacour et al., 2015). The soil organic carbon is simulated by a CENTURY-type model (Parton et al., 1987) and is partitioned in three pools (slow, passive, active) with different residence times.

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022



Model Set-up

The set-up of the simulations performed with ORCHIDEE depends on the data assimilated. The model is run at site scale for the assimilation of eddy-covariance measurements, at spatial resolution 0.72° for the assimilation of the satellite NDVI data, and at the resolution of the atmospheric transport model LMDz (3.75°x2.5°) for the assimilation of atmospheric $CO_2$ measurements. The Olson land cover classification at 5 km is used to derive the PFT fractions at each spatial resolution, but for the flux tower simulations where the proportion of each PFT is set based on expert knowledge. For satellite pixels and global simulations, ORCHIDEE is forced using the 3-hourly ERA-Interim gridded meteorological forcing fields (Dee et al., 2011) (aggregated at 3.75°x2.5° when assimilating atmospheric $CO_2$ concentrations). For the flux tower simulations, the model is forced by local measurements of the meteorological variables at a half-hourly time step.

For each spatial resolution, a prior spin-up simulation was performed by recycling available forcing data. The objective was to bring the different soil carbon reservoirs to “realistic” values, albeit the spin-up runs result in neutral net carbon flux by construction. Each spin-up simulation was then followed by a transient simulation (starting from the first year of measurement for each data stream) and accounting for the secular increase of atmospheric $CO_2$ concentrations.

### ***2.1.2 LMDz***

Model description

The study relies on version 3 of the Atmospheric General Circulation Model of the Laboratoire de Météorologie Dynamique (LMDz) (Hourdin et al., 2006) as implemented for the IPSL contribution to CMIP4. In order to save computational time, we used LMDz in the form of a precomputed Jacobian matrix at a set of $CO_2$ measurement stations (§2.2.3) (see details in Peylin et al., 2016).

Model set-up

To simulate atmospheric $CO_2$ concentrations that can be compared to observations, the transport model has to be forced not only by terrestrial biospheric fluxes (calculated by ORCHIDEE), but also by other natural (e.g. ocean) and anthropogenic $CO_2$ fluxes. We imposed a net emission due to land use change (i.e. deforestation) of 1.1 $GtC.yr^{-1}$ although we also accounted for a larger flux from biomass burning but compensated partly by forest regrowth (see Peylin et al. (2016) for more details). The global maps of biomass burning emissions were taken from the Global Fire Emission Database version 3 dataset (Van der Werf et al., 2006; Randersen et al., 2013) over the period 1997-2010 at a monthly time step and gridded at 0.5°x0.5° resolution. The global fossil fuel $CO_2$ emission products

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022



used here were developed by University of Stuttgart/IER based on EDGAR v4.2 and were provided at a 0.1°x0.1° spatial resolution and at a monthly time scale. The ocean flux component was obtained from a data-driven statistical model based on artificial neural networks that estimated the spatial and temporal variations of the air-sea $CO_2$ fluxes (Peylin et al., 2016).

## 2.2 Assimilated data

### ***2.2.1 in situ flux measurements (F)***

The NEE and LE measurements come from the FLUXNET global network. We used harmonized, quality-checked and gap-filled data (Level 4) at 68 sites from the La Thuile global synthesis dataset (Papale, 2006). The site location is presented in Figure 1. These ecosystem measurements cover very different time spans, ranging from one single year at some sites up to nine years. They constrain seven PFTs among the twelve natural vegetation types represented in ORCHIDEE: tropical evergreen broadleaf forest – TrEBF (3 sites corresponding to 6 site-years), temperate evergreen needleleaf forest – TeENF (16 sites, 45 sites-years), temperate evergreen broadleaf forest – TeEBF (2 sites, 4 site-years), temperate deciduous broadleaf forest – TeDBF (11 sites, 37 site-years), boreal evergreen needleleaf forest – BoENF (12 sites, 44 site-years), boreal deciduous broadleaf forest – BoDBF (3 sites, 6 site-years), and C3 grassland – C3GRA (21 sites, 56 site-years). We assimilated daily-mean values of NEE and LE observations, but only when at least 80% of the 48 potential half-hourly data in a day are available.

### ***2.2.2 Satellite products (VI)***

The NDVI products considered here are derived from MODIS collection 5 surface reflectance data acquired in the red and near-infrared channels and corrected from the directional effects (Vermote et al. (2009). The daily data at 0.72° spanning the period 2000-2010 already assimilated into ORCHIDEE and described in MacBean et al. (2015) are considered. Five among the six deciduous, non-agricultural, PFTs of ORCHIDEE were optimized in this study: TrDBF - tropical broadleaved rainy green forest, TeDBF, BoDBF, BoDNF – Boreal needleleaf summergreen forest, and C3GRA. C4 grasses and evergreen PFTs were not considered. For each PFT, fifteen 0.72° pixels were selected for assimilation depending on their thematic homogeneity with respect to the considered PFT (fractional coverage above 60%) and consistency between the observed NDVI time series and the prior ORCHIDEE. The location of these satellite pixels is shown in Figure 1.

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022



#### 2.2.3 Atmospheric $CO_2$ measurements (CO2)

The surface atmospheric $CO_2$ concentration data come from three databases: The NOAA Earth System Laboratory (ESRL) archive (ftp://ftp.cmdl.noaa.gov/ccg/co2/), the CarboEurope IP project (http://ceatmosphere.lsce.ipsl.fr/database/index_database.html), and the World Data Centre for Greenhouse Gases of the World Meteorological Organization Global Atmospheric Watch Programme (http://gaw.kishou.go.jp). The data include *in situ* measurements, made by automated quasi-continuous analysers, and air samples collected in flasks and later analyzed at central facilities. In this study, we used monthly-mean values of these measurements (Peylin et al., 2016). Ten years of observations over the 2000-2009 period were used from a total of 53 stations located around the world (Figure 1).

### 2.3 Assimilation methodology

#### 2.3.1 Data assimilation framework

The ORCHIDEE model data assimilation system (ORCHIDAS) has been described in previous studies regarding the assimilation of these data streams alone (Kuppel et al., 2012; Santaren et al., 2014; MacBean et al., 2015; Bastrikov et al., 2018) or their combinations (Bacour et al., 2015; Peylin et al., 2016). The assimilation system relies on a Bayesian framework that optimizes ORCHIDEE parameters gathered in a vector **x**, by finding the minimum of a global misfit function $J(\mathbf{x})$. $J(\mathbf{x})$ is a linear combination of the misfit functions associated with each data stream. It is assumed that the errors of observations and on the model parameters are Gaussian and that the data streams errors are independent from each other:

$$
\begin{aligned}
J(\boldsymbol{x}) \quad = \quad & \frac{1}{2}\big[\left(H_{LMDz}\circ H_{ORCH}(\boldsymbol{x}) - \boldsymbol{y}^{\boldsymbol{CO2}}\right)^{t}.\mathbf{R}_{\boldsymbol{CO2}}^{-1}.\left(H_{LMDz}\circ H_{ORCH}(\boldsymbol{x}) - \boldsymbol{y}^{\boldsymbol{CO2}}\right) + \\
& \left(H_{ORCH}(\boldsymbol{x}) - \boldsymbol{y}^{\boldsymbol{F}}\right)^{t}.\mathbf{R}_{\boldsymbol{F}}^{-1}.\left(H_{ORCH}(\boldsymbol{x}) - \boldsymbol{y}^{\boldsymbol{F}}\right) + \\
& \left(H_{ORCH}(\boldsymbol{x} - \boldsymbol{y}^{\boldsymbol{VI}})\right)^{t}.\mathbf{R}_{\boldsymbol{VI}}^{-1}.\left(H_{ORCH}(\boldsymbol{x}) - \boldsymbol{y}^{\boldsymbol{VI}}\right) + \\
& \chi^{t}.\chi\big]
\end{aligned} \tag{1}
$$

where $\boldsymbol{y}^{o}$ are the observation vectors (with $o$ = *F (flux)*, *VI (satellite NDVI)*, or $CO_2$ (concentration); $H_{ORCH}$ and $H_{LMDz}$ are the observational operators of the ORCHIDEE and LMDz models, respectively. $\mathbf{R}^{o}$ is the error covariance matrix characterizing the observation errors with respect to the model (therefore including the uncertainty in the model structure) associated to data stream $o$. The dimensionless control vector $\chi$ quantifies the distance between the values of the optimized parameters and the corresponding prior information $\boldsymbol{x}^{b}$: $\chi = \boldsymbol{B}^{-1/2}.(\boldsymbol{x} - \boldsymbol{x}^{\boldsymbol{b}})$, where **B** is the associated *a priori* error covariance matrix.

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022




We use the gradient-based L-BFGS-B algorithm (Byrd et al., 1995; Zhu et al., 1997) to minimize $J(\boldsymbol{x})$ iteratively. It accounts for bounds in the parameter variations. The algorithm requires the gradient of the misfit function as an input in order to explore the parameter space:

$$\begin{aligned}\nabla_x J(\boldsymbol{x}) \quad = \quad & \mathbf{H_{ORCH}^{CO2}}^{t}.\mathbf{H_{LMDz}}^{t}.\mathbf{R}_{\boldsymbol{CO2}}^{-1}.(H_{LMDz}\circ H_{ORCH}(\boldsymbol{x})-\boldsymbol{y}^{\boldsymbol{CO2}})+ \\ & \mathbf{H_{ORCH}^{F}}^{t}.\mathbf{R}_{\boldsymbol{F}}^{-1}.(H_{ORCH}(\boldsymbol{x})-\boldsymbol{y}^{\boldsymbol{F}})+ \\ & \mathbf{H_{ORCH}^{S}}^{t}.\mathbf{R}_{VI}^{-1}.(H_{ORCH}(\boldsymbol{x})-\boldsymbol{y}^{\boldsymbol{VI}})+ \\ & \mathbf{B}^{-1/2}.\chi\end{aligned} \tag{2}$$

The calculation of uses the Jacobian matrix of ORCHIDEE associated to each data stream, $\boldsymbol{H_{ORCH}^{o}}$ (assuming local linearity of the model), and that of LMDz. For most of ORCHIDEE parameters, is calculated thanks to the tangent linear model of ORCHIDEE obtained by automatic differentiation using the TAF tool (Giering et al., 2005); however, for a few parameters involved in threshold conditions of the model processes, especially related to phenology, we use a finite difference method.

After optimization, the posterior error covariance matrix **A** (for "analysis") of the optimized parameters can be calculated as a function of the Jacobian matrix associated to the gradients of the model outputs with respect to the parameters at the solution for each data stream:

$$\mathbf{A} = \left[\sum \mathbf{H_o^{\infty}}^{t}.\mathbf{R_o^{-1}}.\mathbf{H_o^{\infty}} + \mathbf{B^{-1}}\right]^{\mathbf{-1}} \tag{3}$$

It is computed under the hypothesis of model linearity in the vicinity of the solution. The square root of the diagonal elements of **B** or **A** correspond to the standard deviation $\sigma$ on model parameters.

### ***2.3.2 Parameters to be optimized***

We chose to optimize a limited set of carbon-cycle related parameters of ORCHIDEE as a result of preliminary sensitivity analyses and past DA studies. A short definition of these parameters that mostly control photosynthesis, phenology and respiration, is provided in Table 1, while their associated prior values, bounds and uncertainty are documented in Supplementary Table S3. More comprehensive descriptions of their role in the model processes are provided in Kuppel et al. (2012) and MacBean et al. (2015). The size of soil carbon pools drives the magnitude of the net carbon fluxes exchanged with the atmosphere to a large extent and is closely related to the land use history.

Given that we do not have access to that information, neither at the site scale (for assimilation of NEE measurements) nor at the global scale (for assimilation of atmospheric $CO_2$ concentrations), we use a steady state assumption where ORCHIDEE has been brought to near equilibrium with a long spin-up of the soil carbon pools. To correct for this bias, the initial state of the soil carbon reservoirs is optimized using a multiplicative parameter of both the slow and passive pools as in Peylin et al.(2016). Two multiplicative parameters are used depending on the type of data considered (and their associated spatial scale): for *in situ* flux measurements, we considered site-specific parameters $K_{soilC,site}$; for atmospheric $CO_2$ concentration data, instead of resolving the initial conditions for all grid cells we scaled the carbon pools for 30 large scale regions $K_{soilC,reg}$. Note that having correct soil carbon pools is less important when assimilating satellite NDVI data because these are more closely related to carbon uptake rather than net carbon flux. In total, up to 182 parameters are optimized depending on the data streams considered.

The prior values $\boldsymbol{x}^b$ of the parameters are set to the standard values of ORCHIDEE (Supplementary Table S3). Not all parameters are constrained by all three data streams. In particular, satellite FAPAR/NDVI products inform the timing of phenology of plant vegetation (start and end of the growing season) rather than on photosynthesis or respiration with our DA system (Bacour et al., 2015; MacBean et al., 2015). The dependency of each parameter with respect to the assimilated data streams is indicated in Table 1.

#### ***2.3.3 Data assimilation experiments***

Different data assimilation experiments were tested in order to understand the respective constraint brought by each data stream and evaluate their compatibility with each other and with the model. First, each data stream was assimilated separately and then its combinations with the other two were considered. Second, the three data streams are assimilated altogether. The various experiments are described in Table 2 with the number of data points assimilated and the number of parameters optimized. Indeed, the number of optimized parameters differs with the type of data assimilated as described in §3.2 and in Table 1. The assimilations have a high computational cost, with an average value for joint assimilations using all three data streams of about 50,000 hr Central Processing Unit time on AMD Rome compute nodes at 2.6 GHz with 256 GB memory per node.

Two assimilation experiments combining the three data streams were tested: one experiment (F+VI+CO2) with all parameters optimized in a single step; and an additional experiment following a 2-step optimization (F+VI+CO2-2steps), as described hereafter. In the first step, the global soil carbon reservoirs are constrained by assimilating atmospheric $CO_2$ data only, and optimizing the two main

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022




parameters controlling soil respiration, $KsoilC_{reg}$ and $Q10$. In the second step, all parameters but $KsoilC_{reg}$ were optimized from the three data streams: $KsoilC_{reg}$ was retained from the first step and $Q10$ was optimized but the prior uncertainty for $Q10$ for the second step corresponded to the posterior uncertainty derived from the first step. We did this to correct for the initialisation of the soil carbon pools following model spin-up; the motivations and implications of the two assimilations experiments are further discussed in the result and discussion sections.

The results of these assimilations were compared to the companion study of Peylin et al. (2016) in which the same data streams were assimilated in a sequential/stepwise approach: NDVI data were assimilated first, then *in situ* flux measurements, and finally atmospheric $CO_2$ concentration measurements. In addition, there are a few differences in the set-up compared to the present study, in particular the use of only three years of atmospheric $CO_2$ data (cf. details provided in Supplementary Text S1).

### *2.3.4 Error statistics on observations and parameters*

#### *2.3.4.1 Observation error statistics*

Like in previous studies with ORCHIDAS, we defined **R** as diagonal and computed the variances from the Root Mean Square Difference (RMSD) between the data and the *a priori* ORCHIDEE simulations (*i.e.* performed with the model default parameter values) for fluxes and satellite observations. However, it is worth noting that this approach overestimates the variances in order to compensate for any neglected correlations. For atmospheric $CO_2$ measurements, we followed a different methodology given the large discrepancy in the modelled *a priori* concentrations: the errors were determined as a function of the observed and modelled temporal concentration variability at each site (Peylin et al., 2016), thus neglecting the structural errors of the terrestrial model.

#### *2.3.4.2 Tuning of the prior error statistics*

We assumed that errors in the prior parameter values are independent and therefore we used a diagonal **B** matrix. We populated the diagonal of **B** in an iterative way from consistency diagnostics of the data assimilation system following Desroziers et al. (2005), as described hereafter. If both **B** and **R** matrices are correctly specified and if the estimation problem is linear, they should be related to the covariance of the residuals (**d**) between observations and background simulations (*i.e.* innovation) following:

$$\mathbf{H_o}.\mathbf{B}.\mathbf{H_o}^t + \mathbf{R} = E\left[(\boldsymbol{y^o} - H(\boldsymbol{x^b})).(\boldsymbol{y^o} - H(\boldsymbol{x^b}))^t\right] = E\left[\mathbf{d_b^o}.\mathbf{d_b^o}^t\right] \quad (4)$$

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022




With

$$\mathbf{R} \ = \ E\left[(\boldsymbol{y^o} - H(\boldsymbol{x^a})).(\boldsymbol{y^o} - H(\boldsymbol{x^b}))^t\right] = E\left[\mathbf{d_a^o}.\mathbf{d_b^o}^t\right] \tag{5}$$

$$\mathbf{H_o}.\mathbf{B}.\mathbf{H_o}^t \ = \ E\left[(H(\boldsymbol{x^a}) - H(\boldsymbol{x^b})).(\boldsymbol{y^o} - H(\boldsymbol{x^b}))^t\right] = E\left[\mathbf{d_b^a}.\mathbf{d_b^o}^t\right] \tag{6}$$

Similarly, the diagnostic on analysis errors can be determined from the residuals between observations and posterior simulations as:

$$\mathbf{H_o}.\mathbf{A}.\mathbf{H_o}^t \ = \ E[(H(\boldsymbol{x^a}) - H(\boldsymbol{x^b})).(\boldsymbol{y^o} - H(\boldsymbol{x^a}))^t] = E\left[\mathbf{d_b^a}.\mathbf{d_a^o}^t\right] \tag{7}$$

In principle, the tuning of **B** and **R** needs to be performed iteratively for successive values of and of the corresponding residuals, until convergence, which is prohibitive in terms of computing time. The estimation of the covariance matrices depends on the mathematical expectation (*E*) which would require several realizations of the residuals to diagnose the error statistics (Desroziers et al. (2005); Cressot et al., 2014). In this study, only one optimization was performed using one set of *a priori* parameters for each dataset. We therefore calculated these metrics by averaging the diagonals of the matrices described by both sides of the equations for all available observations (Kuppel et al., 2013). This way, both sides are scalar values (Cressot et al., 2014).

The standard deviation of the errors were determined after a few trials considering the three single data stream assimilation experiments independently: For each DA experiment we started from an initial parameter error set at 40% of the variation interval for each parameter (as in Peylin et al., 2016); The errors were then varied in order to fulfill the consistency diagnostics on the parameter and observation errors (see Supplementary Text S3). Finally, we evaluated the consistency of the resulting model-data covariance matrices for the DA experiments with multiple data streams using the reduced chi-square test (i.e. the chi-square statistic normalized by the number of observations, *m* (Chevallier et al., 2007; Klonecki et al., 2012), which is implicitly optimized by the Desroziers et al. (2005) approach:

$$\chi^2 = \frac{2J(\boldsymbol{x^a})}{m} \tag{8}$$

If the **R** and **B** covariance matrices are well defined, the ratio of each term of the diagnostics of Desroziers et al. (2005) (ratio between **R** and $E\left[\mathbf{d_a^o}.\mathbf{d_b^o}^t\right]$ ; and $E\left[\mathbf{d_b^a}.\mathbf{d_b^o}^t\right]$ ; and and

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022




$E\left[\mathbf{d_b^o}.\mathbf{d_b^o}^t\right]$) should approach 1. Table 3 shows the values of the consistency diagnostics for the final parameter error set-up.

The diagnostics for **R** (ratios slightly above 1 for all data streams**)** and for the reduced chi-square (Table S1 - values below 1) indicates a slight overestimation of the observation error. The diagnostics for **B** ($ratio^B$) show a stronger overestimation of the *a priori* error for NEE, LE and atmospheric $CO_2$, but an underestimation for NDVI. For fluxes and satellite data, the combined diagnostics for **R** and **B** ($ratio^{BR}$) appear consistent with ratios close to 1. For CO2 however, the value of $ratio^{BR}$ close to the value of $ratio^B$ highlights the strong influence of the background information (**B** matrix) or the model structure on the optimization, while the large value of expresses a strong underestimation of the observation error. Indeed, when determining $\mathbf{R_{CO2}}$, we purposely did not account for the large bias (by about 1 ppm.yr$^{-1}$) between the observed $CO_2$ temporal profiles at stations and the prior simulations, which is due to the initialisation of ORCHIDEE's carbon pools (which is discussed in the Result section).

Finally, for the diagnostics on the analysis, the various tests performed (Supplementary Text S3) all lead to negative quantities. Instead, the simulations of the calibrated model were expected to be contained in between their prior state and the observations (the residuals having opposite signs, their product is positive). This result may reflect a too strong model correction. However, it should be noted that a strong assumption associated with these tests concerns the linearity of the model, which may not hold for terrestrial biosphere models.

## 2.4 Diagnostics for system evaluation

### ***2.4.1 <u>Optimisation performance</u>***

We measured the efficiency of any assimilation by quantifying the reduction of the cost function as the ratio of the prior to posterior values. It should be noted that the minimum value of the cost function is not expected to be zero given the uncertainty in both the data and model, and the limited number of degrees of freedom (number of optimized parameters) allowed. We also looked at the ratio of the norm of the gradient between the prior and posterior misfit functions, as it illustrates the progression towards the expected optimum, for which the gradient is null. The decrease of the norm of the gradient depends on the estimation problem (non-linearities, number of observations versus number of optimized parameters, constraints of the data on the model processes, etc.); however, based on our experience with non-linear problems, we still expect the norm of the gradient to be reduced by at least two orders of magnitude.

The analysis of the optimization performances are summarized in §3.1 and detailed in Supplementary Text S4.

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022



### 2.4.2 Model improvement and posterior predictive checks

The model improvement was quantified by the reduction of the root mean square deviation (RMSD) between model and data, prior and posterior to optimization, expressed in %, as $100 \times (1 - RMSD_{post}/RMSD_{prior})$.

We conducted posterior predictive checks by running the model optimized after assimilation of one or two data streams and quantifying the resulting model-improvement with respect to the data streams not accounted for in the assimilation.

### 2.4.3 Uncertainty reduction on parameters and error budget

The knowledge improvement on the model parameters brought by assimilation was assessed by the uncertainty reduction determined by 1- $\sigma_{post}/\sigma_{prior}$, where $\sigma_{post}$ and $\sigma_{prior}$ are the standard deviation derived from the posterior (**A**) and prior (**B**) covariance matrices on the model parameters and output variables.

A comprehensive quantification of the uncertainty reduction on model variables would require accounting also for the covariance matrix of the model structural error which could be the dominant factor. Because this covariance matrix is difficult to estimate (see Kuppel et al., 2013, for a first attempt in the case of the NEE), we instead analyzed the posterior errors on NEE and GPP at regional to global scales, as the projection of the posterior error on parameters in the space of the model variables. The posterior error on C fluxes is then characterized by the covariance matrix $\mathbf{R^a}$ as:

$$\mathbf{R^a} = \mathbf{H_o^\infty} . \mathbf{A} . \mathbf{H_o^\infty}^t \quad (9)$$

with the Jacobian matrix $\mathbf{H_o^\infty}$, being the first derivative of the target quantity (e. g., NEE, GPP) to the optimized parameters derived from an assimilation experiment *o*.

### 2.4.4 Assessment of the information content of each data stream

For the joint assimilations using the three different data streams, we further analyzed the influence matrix **S** that quantifies their leverage on the model-data fit (Cardinali et al., 2004):

$$\mathbf{S} = \mathbf{R^{-1}} . \mathbf{H^\infty} . \mathbf{A} . \mathbf{H^\infty}^t \quad (10)$$

A diagonal element $S_{ii}$ is the rate of change of the simulated observable *i* with respect to variations in the corresponding assimilated observation *i*. $S_{ii}$ is referred to as "self-sensitivity" of "self-influence". A zero self-sensitivity indicates that this $i^{th}$ observation does not contribute to improving its simulation

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022




by the model, whilst $S_{ii}$ = 1 indicates that the fit of the sole observation *i* mobilizes an entire degree of freedom (*i.e.* one parameter). In addition to the total influence matrix (equation 10), we also determined the partial influence matrices associated to each data stream *o*, using the corresponding diagonal $\mathbf{R_o}$ matrices and $\mathbf{H}^{\infty}$ in equation 10.

We analyzed the trace (i.e. the sum of all diagonal elements) of **S** that quantifies a measure of the amount of information that can be extracted from all observations / all data streams. We used two derived quantities: the global average observation influence (OI) and the relative degrees of freedom for signal (DFS) associated with the data stream *o,* which measures its relative contribution to the fit. They are defined as follow (with *m* the total number of observations):

$$OI = \frac{tr(\boldsymbol{S})}{m} \tag{11}$$

and

$$DFS = 100 \times \frac{tr(\boldsymbol{S_o})}{tr(\boldsymbol{S})} \tag{12}$$

# 3 Results and discussion

## 3.1 Model improvement for the different assimilation experiments

### 3.1.1 Cost function reduction

The reduction of the cost function varies between the different experiments with the lowest reductions for the single data streams experiments F and VI (around 10%). However, the correction of the model-data misfit when $CO_2$ data are assimilated is much higher (at least factor of 10 reduction). Noteworthy, this strong model improvement is obtained for a lower departure of the parameters from their prior values than when fluxes or satellite data are assimilated (cf. section 3.3, and Figure 6).

A detailed description of the optimization performances with respect to the minimisation of the cost function is detailed in Supplementary Text S4 and Table S2.

### 3.1.2 Overall fit to the observations

The impact of assimilating one type of observation on all the data streams (including those that are not assimilated) was evaluated for the various assimilation experiments. The reduction of the model-data mismatch (i.e. reduction in prior RMSD) after assimilation of each data stream (or any combination of them) is illustrated in Figure 2. The length of the boxes (first and third quartiles) of the whisker plots highlight the spread in misfit reduction across sites/vegetation types. For fluxes,

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022




only the impact on NEE is shown, given the choice of optimizing parameters is mostly related to the carbon cycle. Using the parameter values optimized in either the F and VI assimilations has a strong detrimental impact on the simulated atmospheric $CO_2$ data because the soil carbon pools were not adjusted in these DA experiments. Therefore, we also analyzed the changes induced on the detrended seasonal cycles of atmospheric $CO_2$ concentrations (hence removing the trend using the time series decomposition based on the CCGCRV routine (Thoning et al., 1989), as described in Supplementary Text S2) (Figure 2c).

For a given data stream, the improvement is usually better for the experiment where that data stream is assimilated alone (compared to joint assimilation of two or more data streams; Figure 2). One noteworthy exception is the assimilation of NDVI alone (VI experiment) that results in a lower model improvement with respect to NDVI than when it is assimilated in combination with other data-streams. Only the phenology parameters are optimized with NDVI alone. In contrast, a higher number of parameters are optimized in the joint assimilations involving NDVI, which hence improves the timing of phenology and the amplitude of the annual cycle when flux or atmospheric CO2 data are also assimilated. For both experiments F and VI, the reduction of the model-data misfit can be negative, which reflects how the assimilation can degrade the model performance for a few pixels/sites by searching for a common parameter set. This is not observed with the assimilation of atmospheric $CO_2$ data only for which the optimized model is always closer to the observations than the prior model (due to a correction of the $CO_2$ trend), at all stations (see Supplementary Text S5 for a detailed description of the reduction in model-data misfit each single-data stream assimilation experiment (F, VI, CO2)).

The collateral impact of assimilating one data stream on the other simulated observables is evident in the misfit reductions shown in Figure 2 (e.g., examine the "VI" experiment on the NEE misfit reduction in Figure 2a). While using optimized phenological parameters retrieved from satellite data alone (experiment VI) degrades the modelled seasonality of NEE as compared to the measurements (median RMSD reduction of -3%), the optimization with respect to *in situ* flux data (F), with additional control parameters, leads to a general improved consistency between modelled FAPAR and satellite NDVI time series (median RMSD reduction of 8%). The impact on LE is much lower for all DA experiments (median values close to 0% in all cases, result not shown). One can also note the positive impact of the F and VI assimilations on the atmospheric $CO_2$ data with median RMSD reductions of 15.8% and 11.2% respectively for the detrended time series. Such an improvement after assimilation of *in situ* flux data corroborates the findings of Kuppel et al. (2014) and Peylin et al. (2016). Noteworthy, this improvement is of the same order as that achieved when assimilating

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022


atmospheric $CO_2$ data alone (median RMSD reduction of 14%). The parameters retrieved from the CO2 experiment have also a small but positive impact at the site level with respect to NEE (median value of 3%) and FAPAR (0.8%). These low values are explained by the fact that, in the CO2 assimilation, most of the model improvement is attributable to the tuning of the soil carbon pools (to fit the atmospheric CO2 growth rate) while the other parameters are marginally changed (see Figure 3).

For the joint assimilation experiment (F+VI, F+CO2, VI+CO2, or F+VI+CO2; Figure 2), the model-data agreement is improved for all assimilated data streams, as expected, while the model degradation relative to the data not assimilated is generally not as severe as compared to the assimilation of individual data stream experiments described above, with the exception of the F+VI experiment. The simultaneous assimilation of flux measurements and satellite NDVI data (F+VI) leads to enhanced model improvement compared to when these data are assimilated alone (cf. Supplementary Text S5). In the simultaneous assimilations involving atmospheric $CO_2$ data, most of the model improvement concerns CO2 (Figure 2c) while the benefit for the fluxes and FAPAR/NDVI is weak (RMSD reduction below 3%). Noteworthy, the 2-step assimilation F+VI+CO2 (see Section 2.3.3) results in an even higher model improvement for both NEE and FAPAR than the 1-step approach.

The misfit reduction for the raw (i.e., not detrended) atmospheric $CO_2$ data is high (median reduction ~75%) and remains quite stable among the various different combinations of data streams that include atmospheric CO2 (Figure 2c solid bars experiments including “CO2”), with the exception of the F+VI+CO2-2steps experiments. The misfit reductions for the detrended $CO_2$ time series are generally lower (median reduction less than ~15%) and there are more pronounced differences between experiments. Again, these results highlight the predominance of the correction of the trend in atmospheric $CO_2$ time series through the fitting of the carbon pool parameters, over the tuning of the other model parameters related to photosynthesis and phenology. The 2-step approach permits to partially overcome that limitation, with the improvement of the mean seasonal cycle for the three data streams (Figure 2c).

#### 3.1.3 Specific improvements at $CO_2$ stations

Figure 3 further analyzes the impact of each assimilation experiment on the fit to the observed atmospheric $CO_2$ concentrations in terms of the bias in the long-term trend (2000-2009) and fit to the

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022



mean seasonal cycle over the same period (i.e., bias in seasonal amplitude and length of the carbon uptake period). For the trend analysis (Figure 3a), only experiments where atmospheric $CO_2$ measurements are assimilated are considered.

With the default parameter values, the simulated fluxes by ORCHIDEE lead with $LMD_Z$ to overestimates the (trend) by about 1 ppm.yr$^{-1}$. When assimilating atmospheric $CO_2$ data, most of the parameter correction aims at reducing this bias. This is mostly achieved by tuning the regional $K_{soilC\_reg}$ parameters: the net land carbon sink is increased globally in order to match the observed trend at most stations (reducing the bias from around 1 ppm.yr$^{-1}$ to 0.1 ppm.yr$^{-1}$). Compared to the improvement in the bias in the trend, the improvements (reduction in bias) in the amplitude of the $CO_2$ seasonal cycle and in the CUP length (Figures 3b and c) are marginal. Note that our joint DA experiments lead to significantly lower trend biases compared to the stepwise approach, which is probably due to the longer period of the atmospheric $CO_2$ data considered (10 years vs 3 years for the stepwise).

For the amplitude of $CO_2$ concentrations, the joint assimilations including $CO_2$ data lead to lower improvements on average compared to any single data stream assimilation experiment. Interestingly, the highest improvements in $CO_2$ amplitude are achieved when flux data are assimilated (F or F+VI), which reveals that the constraint on photosynthesis and respiration provided by FLUXNET measurements is consistent with the amplitude of the seasonal atmospheric $CO_2$ cycle and within the ORCHIDEE-LMDz model (as already pointed out in Kuppel et al. (2014)). Surprisingly, the use of satellite vegetation indices (VI) leads to a slightly lower residual amplitude bias than when atmospheric $CO_2$ data are assimilated, albeit a lower number of optimized parameters. For the length of the carbon uptake period (CUP), the relative model correction appears small for almost all experiments and is lower than what is achieved for the trend and amplitude. Some degradation (increased model-data bias) is even obtained for the cases F and F+CO2. This may be attributed to some inconsistency in the phasing of the CUP deduced from the FLUXNET stations and from the atmospheric stations (given differences in the spatial and temporal scale constraints brought each data stream). Among the single data stream assimilations, the highest improvement is obtained for VI where the optimisation of the phenological parameters was the only improvement allowed for tuning the model. For the joint assimilations, those combining the three data streams provide the best performance and perform better than the stepwise approach.

Among the joint assimilations with three data streams, the 2-step approach results in the largest reduction in amplitude and CUP bias, but, on the other hand, the larger trend bias .

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022




## 3.2 Impact of the assimilations on regional to global land C fluxes and errors

Figure 4 now compares the carbon fluxes (NEE and GPP) at the global scale and for three large regions (northern and southern extra-tropics, and tropics) using hindcast simulations based on the different optimisations.

NEE is close to equilibrium by construction in the prior model (about -0.3 GtC.yr$^{-1}$ globally). Note first that experiments excluding $CO_2$ data produce land carbon fluxes (from -10 (F+VI) to +6 (VI) GtC.yr$^{-1}$, not shown in Figure 3) that are not compatible with our understanding of the land C fluxes. These were however obtained without a new spin-up nor any transient simulations like it is done for the TRENDY experiment (Sitch et al., 2015). For all experiments including atmospheric $CO_2$ data, the assimilations lead to much more negative NEE (increased land carbon sink) compared to the prior for nearly all regions. For these experiments that include $CO_2$ data, the optimized carbon sinks are about -2.4 GtC.yr$^{-1}$ at the global scale, with the exception of the stepwise approach, which is -1.7GtC.yr$^{-1}$ (see Supplementary Text S6 for detailed results for each assimilation experiment). Therefore, our joint assimilation with atmospheric $CO_2$ data results in a land C sink that is in the range of independent TBM estimates of the global net carbon budget (over the same period, the Global Carbon Project reports a global land sink of -2.9 GtC.yr$^{-1}$ ± 0.8 standard deviation (see Table 5 of Friedlingstein et al., 2020). Note that we have imposed (see method in §2.1.2) a net emission from land use change (i.e. deforestation) of +1.1 GtC.yr$^{-1}$ (2001-2009) which is slightly lower than that reported in Friedlingstein et al. (2020) from the TBMs (1.6±0.5 GtC.y$^{r-1}$) or the Bookkeeping methods (1.4±0.7 GtC.yr$^{-1}$), hence our lower terrestrial carbon sink. Note that the lower terrestrial sink obtained with the stepwise approach is partly due to the optimisation against atmospheric $CO_2$ data over a different and shorter period (2001-2004) with a lower C sink. Using a three year period is also likely to be not enough for constrain the ORCHIDEE model due to large year to year variability of the terrestrial C sink.

These similar posterior global scale budgets however hide significant regional contrasts. While the three joint assimilation experiments F+CO2, VI+CO2, and F+VI+CO2, lead to similar NEE budgets across regions, the CO2 and F+VI+CO2-2steps experiments result in distinctly different estimates. In the northern extra-tropics, the CO2 assimilation results in the largest C sinks (numbers provided in Supplementary Text S6) while the F+VI+CO2-2steps assimilation leads to the lowest C sink, with a magnitude that matches the stepwise assimilation set-up (Peylin et al., 2016). The reverse is obtained for the Tropics.

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022




Partitioning the land carbon budget between the northern extra-tropics and the tropics is a key scientific question, in particular because of differences in the C residence time between these ecosystems. From the global carbon budget, a discrepancy exists between the partition estimated by the atmospheric $CO_2$ inversions and by the terrestrial biosphere models. Atmospheric inversions estimate a larger sink over the northern extra-tropics than TBMs (around 1.8 GtgC.yr$^{-1}$ versus 1.0 GtC.yr$^{-1}$ for the period 2010-2020), although with large variations between TBMs (Friedlingstein et al., 2020, Figure 8). Conversely, TBMs estimate a larger C sink over the tropics (Ahlström et al., 2015; Sitch et al., 2015) than the inversions, which estimate an approximately net neutral C sink (Peiro et al., 2022). The stepwise and F+VI+CO2-2steps assimilations follow the typical partitioning pattern of TBMs' behavior, with a stronger C sink in the tropics than in the northern hemisphere. On the opposite, the three two or more data stream experiments F+CO2, VI+CO2 and F+VI+CO2 lead to an approximately equal C sink in the northern hemisphere and tropics (thus unlike the general pattern for TBMs). In contrast, the CO2 experiment leads to a similar regional partitioning as the atmospheric inversions. For the F+VI+CO2-2steps experiment, the tropical sink is almost doubled as compared to the other simultaneous assimilation experiments in spite of a slightly reduced GPP. The correction of the $CO_2$ trend in the first step of the 2-step approach, with the optimisation of the soil disequilibrium ($K_{soilC\_reg}$ parameters), tends to favor a tropical C sink, which has a direct impact on the atmospheric $CO_2$ stations of both hemispheres. Although such an approach offers the advantage of correcting more substantially the seasonal cycle of atmospheric $CO_2$ (see section 3.1.3), it may still be sub-optimal for partition of the terrestrial C sink (between northern and tropical regions).

With a global scale budget of 171 GtC.yr$^{-1}$ for GPP, the prior ORCHIDEE model is on the high range of recent estimates of the global GPP, as synthesized in Anav et al. (2015), the mean value of which being around 140 GtC.yr$^{-1}$. Depending on the data assimilated in this study (note that we include all assimilation experiments in estimating the posterior GPP), the posterior GPP ranges from 147 GtC.yr$^{-1}$ (F+VI) to 170 GtC.yr$^{-1}$ (VI+CO2) at the global scale. The greatest differences with the prior are obtained for the experiments involving flux and satellite data (alone or the two combined). This is directly linked to large corrections in photosynthesis parameters for these experiments (see §3.3). In comparison, the assimilations involving atmospheric $CO_2$ concentrations data are more conservative with respect to GPP. Assimilating atmospheric $CO_2$ data alone lessens the GPP reduction by a factor of about three compared to assimilations with F and VI data, and the corrections for the joint assimilations using $CO_2$ data is even lower (cf Supplementary Text S6 for details).

By propagating the error on the parameters (see § 3.3) in the observation space (see Eq. 9), we calculated the uncertainty in NEE and GPP fluxes caused by parameter uncertainty for the prior and

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022



optimized models. The error statistics, initially calculated at monthly/grid scale resolutions, were aggregated over the same regions as above, fully accounting for the spatio-temporal correlations between grid cells (Figure 5).

At the global scale, the prior error standard deviation for NEE (4.7 $GtC.yr^{-1}$) is high compared to the typical uncertainty associated to TBMs (about 0.5 $GtC.yr^{-1}$, Friedlingstein et al. (2020)) or to atmospheric inversions (estimated uncertainty ~0.4 $GtC.yr^{-1}$ in Peylin et al.(2013)). This is a consequence of neglecting negative error correlations between them (as done in nearly all C cycle DA studies). Given this high prior uncertainty, the posterior error for NEE and GPP are significantly reduced, as expected. Because of the strong dependence of the posterior errors on the optimisation set-up and the fact we do not consider the error of the model, we should only compare the relative error reduction between DA experiments. Noteworthy, the posterior errors in global NEE obtained for the experiments CO2 and VI+CO2 are about 15 times lower than the posterior errors resulting from the other data combinations (and three orders of magnitude lower than the prior error). This is due both i) to the need for the DA system to correct the large *a priori* mismatch of the atmospheric $CO_2$ growth rate and ii) to the lower number of optimized parameters in these configurations (Table 2: about 60% more parameters being optimized in F+VI+CO2 than in CO2 or VI+CO2). The joint assimilations result in higher posterior errors on NEE, while they usually lead to the lower posterior errors on GPP. For GPP, the lowest posterior errors are found for the experiments combining F and CO2 data, while experiments F, CO2 and VI+CO2 lead to larger posterior errors . This is due to the fact that i) F and CO2 data provide a stronger constraint on the annual mean photosynthesis than VI data and that ii) F and CO2 data provide cross constraints on photosynthesis. Experiment VI, in which about ten times fewer parameters are optimized and targeting primarily the timing of phenology, results in the highest posterior GPP errors (although still a reduction from the prior).

Finally, one can observe that the posterior errors are higher in the tropics for both NEE and GPP (and the reduction compared to the prior error is lower), which is even more prominent in the experiments using *in situ* flux data alone or with satellite data, a direct consequence of the lower data availability (eddy-covariance measurements) to constrain the model parameters for tropical PFTs.

### 3.3 Parameter estimates and associated uncertainties

Figure 6 shows the impacts of the different assimilation experiments on a subset of the retrieved parameter values and their associated uncertainties (the remaining parameters are shown in Figure

S1). The estimates are compared to the posterior parameters from the stepwise approach used in Peylin et al. (2016).

While the stepwise study showed only few changes in the parameter estimates between the sequential steps (and hence as a function of the data stream from which the parameters were constrained) (Peylin et al., 2016), our results show a large variability between the assimilation experiments (either between experiments considering a single data stream or between single- vs joint- assimilations). For most parameters, the highest departures from the prior values are obtained for single-data stream assimilations. Higher changes are obtained for flux or satellite data as compared to the estimates retrieved with atmospheric $CO_2$ data alone which remain closer to the prior values. This reflects the lower constraint brought by the CO2 assimilation experiment on photosynthesis and phenology related processes. This is largely due to the correction of the trend bias via a few respiration related parameters, which prevails over the improvement of the other photosynthesis and phenology parameters.

The joint assimilations (based on two or three data streams) usually result in a lower departure from the background. For the parameters constrained by two data streams, the optimized values generally fall in between those retrieved when these data streams are assimilated alone. This feature shows how the system tries to find a compromise solution and illustrates potential overfitting with only one data stream. The values optimized in the three experiments involving atmospheric $CO_2$ data show little variability for all parameters, except in F+VI+CO2-2steps where the tuning of the multiplicative parameter of regional soil carbon pools $K_{soilC\_reg}$ is decoupled from the optimization of the other photosynthesis and phenological parameters. The decrease of $K_{soilC\_reg}$ parameters from the prior value is very small in all experiments, although these parameters are responsible for most of the correction of the atmospheric $CO_2$ trend. This highlights the challenge of optimizing soil C disequilibrium with our approach based on a model spin-up without a transient run. The smallest $K_{soilC\_reg}$ changes are obtained for the 2-step approach. Note that in this approach, *Q10* is also estimated in the first step; the corresponding estimate is similar to the value retrieved in the second step (which is displayed in Figure 3), below 0.5% difference, and consistent with the estimates of the other joint assimilation experiments. For some parameters/PFTs, the direction of the departure with respect to the prior value (increase or decrease) may differ depending on the data stream assimilated (as detailed in S5).

At the first order, the estimated parameter uncertainties decrease with the number of observations assimilated, as expected from Equation 4, and given that the observations are treated as independent data. However, given that the estimated parameter errors strongly depend on the set-up of **B** and **R** matrices and that we did not use error correlations in these matrices, we should only

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022


focus on the relative error reduction between experiments. The uncertainty reduction achieved through the assimilation of atmospheric $CO_2$ data is usually lower than when flux and satellite data are assimilated alone, and typically vary between 10% and 60% for most photosynthetic and phenological parameters. Most often, the joint assimilations involving two data streams result in an uncertainty reduction higher or of the same order than that achieved in the single-data assimilations. The joint assimilation combining the three data streams generally results in the highest uncertainty reduction, with values typically between 60% and 90%. The values are much higher than those inferred from the stepwise approach (at the last step where three years of atmospheric $CO_2$ data were assimilated), which are more on the order of the uncertainty reduction obtained in the CO2 assimilation experiment.

### 3.4 Relative constraints brought by the different datasets

We now quantify the impact of each of the three data streams on the analysis using the global average observation influence (quantified by OI) and information content (DFS) metrics defined in § 2.4.4. We recall that OI (i.e. trace of **S** normalized by the number of observations) gauges the average influence that each single observation has on the analysis, while the relative DFS measures the overall weight of one data stream in the optimization (the difference between OI and DFS is due to the number of observations assimilated, Cardinali et al. (2014)). OI and DFS are determined for the joint assimilation experiments combining the three data streams.

Because of the very large number of observations (above 300,000) involved in the assimilation, only the diagonal elements of the influence matrix (Eq. 10) can be calculated. The trace of **S** measures the equivalent number of parameters and is equal to 132. Such a value, lower than the number of parameters (182), indicates that the optimized parameters may not be fully independent (although parameter error correlations have been ignored in our **B** matrix) as already reported in Kuppel et al. (2012), or that some are not constrained during the optimisation process (as for instance $LAI_{MAX}$ which estimates remains at its *a priori* value for some PFTs, Figure S1 ).

The values of OI are provided in Table 4 for flux, NDVI and atmospheric $CO_2$ data. With about the same number of observations considered (Table 2, last column), one *in situ* flux measurement has about 10 times more weight than one NDVI observation. This is a consequence of the larger number of parameters constrained by flux measurements than by NDVI data in our set-up. The highest influence is found for atmospheric $CO_2$ data, the relative weight of one atmospheric $CO_2$ measurement being 4 times greater than that of one flux observation, albeit the much lower number of data assimilated. This is a consequence of the strong weight of the mismatch between the *a priori*

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022



simulated and the observed atmospheric $CO_2$ trend (over 10 years), which is drastically reduced through the optimisation.

However, the smaller number of atmospheric $CO_2$ data assimilated, compared to flux and NDVI datasets, reduces the overall constraint on the analysis provided by atmospheric $CO_2$ data, as gauged by its relative DFS. Hence, our optimization is mainly controlled by flux data which have an overall contribution of about 75%, that is about 5 times larger than the constraint brought by atmospheric $CO_2$ data and 7 times larger than that of satellite NDVI. Differences between F+VI+CO2 and F+VI+CO2-2steps are relatively small for both OI and DFS but show a slightly lower weight of atmospheric $CO_2$ data for the 2 steps experiment.

The same analysis was performed by discriminating the influence of each PFT for flux and satellite data, and each station for atmospheric $CO_2$ concentrations (Figure 7 - experiment F+VI+CO2). For the flux data, the results are mainly proportional to the number of observations available (hence, the lower results are obtained for BorDBF, TeDBF and TrEBF, for which the number of assimilated data is about one order of magnitude lower than for the other PFTs; see § 2.2.1).

For satellite NDVI data however, the number of data is the same for each PFT. The discrepancies between PFTs is thus less pronounced than for flux data and related to the ability of the selected parameters to correct the phenology of each PFTs (constrained by the NDVI data). For TrDBF and C3GRA, the inability to correct the start of the growing season ($K_{pheno,crit}$, remains close to the prior values, as seen in Figure 3) may explain the lower contribution of these PFTs.

For atmospheric $CO_2$ data, the DFS is relatively well distributed across stations, with a mean value of 1.9 (range 0.19 – 14.5), in particular in the northern hemisphere. The higher values are found for a few southern hemisphere stations: Halley Station - HBA (6), Syowa - SYO (8.4), South Pole - SPO (11.9) and Cap Grim Observatory - CGO (14.5). Possible reasons for their larger impact may combine: a strong a *priori* model-data mismatch that is substantially corrected, ocean-driven concentration variations not well captured by the prescribed ocean flux but incidentally well corrected by remote land fluxes, etc.

## 4 Summary and outlook

By assimilating separately or simultaneously up to three independent carbon-cycle related data streams (*in situ* measurements of net carbon and latent heat fluxes, satellite derived NDVI data, and measurements of atmospheric $CO_2$ concentration at surface stations) within the ORCHIDEE model (and an offline transport model based on pre-calculated transport fields with LMDz), we have been able to analyze their compatibility, complementarity, and usefulness, in the frame of a global-scale carbon data assimilation system. We investigated how the different combinations of data streams constrain the parameters of the ORCHIDEE land surface model, and by consequence the simulated historical spatial and temporal distribution of the net and gross carbon fluxes (NEE and GPP), as well as FAPAR and atmospheric $CO_2$ concentrations.

Our analyses focussed not only on the model outputs, but also on methodological aspects aiming at 1) checking that the error statistics on parameters and observations were correctly assigned, 2) assessing the optimisation efficiency, and 3) quantifying the relative informational content brought by each data stream. In doing so, the study highlighted some challenges in handling model-data bias in Bayesian optimisation frameworks (in particular the initialisation of the soil carbon pools) and evaluating their impact on the optimized model.

### 4.1 Benefits of simultaneous assimilations

Joint/simultaneous assimilations are more complex to implement compared to stepwise/sequential assimilations. In principle a stepwise approach could lead to similar results than a joint approach, if the posterior parameter error covariance matrix could be fully characterized at each assimilation step and further propagated as prior information in the next step. However, given that this is in practice not feasible, stepwise/joint approaches lead to different optimized models. With a joint assimilation, biases and incompatibilities between data streams may impact more directly a larger set of parameters than in a stepwise assimilation. The characterization of the prior observation errors becomes also more critical as they condition the relative weight of the observations in the misfit function to minimize and their influence on the solution (analysis). Hence, we designed several tests beforehand to refine the configuration of the framework for the simultaneous assimilations, relying on consistency metrics on parameter and observation error statistics of Desroziers et al. (2005). In spite of the limitation of their application to non-linear models like ORCHIDEE, their implementation has proved useful to assign the error statistics on prior parameters relative to the observation error statistics for the different data streams. Ultimately, this has led to an improved consistency of the optimized models at regional and global scales.

Single data stream assimilations usually lead to the best model - data fit for the assimilated data stream, as compared to joint assimilations. However, most often these single data stream assimilations also produce degraded results with respect to the data that were not assimilated. This reveals potential overfitting issues with a higher variability of the optimized parameters than in the joint assimilations. Overfitting is a key issue for DA studies which can be partly alleviated when

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022



combining different data streams within a consistent framework: because they bring different information pieces to the model processes, they contribute to better circumscribing a set of model parameters. Among the several assimilation experiments considered, those where several data were assimilated simultaneously were those resulting in the least model degradation while always showing a model improvement. Joint assimilations regularized the system and resulted in a reduced variability in parameter estimates and in optimized NEE and GPP.

### 4.2 Caveats and perspectives concerning the initialisation of the soil carbon pools

We showed that reaching this global terrestrial carbon sink was mostly achieved by correcting the initial soil carbon reservoirs in the ORCHIDEE model. Their tuning enables the correction of the trend bias between atmospheric $CO_2$ time series measurements at stations and the prior ORCHIDEE-LMDz model. The impact of this trend bias on the optimization performance was highlighted by the quantification of the influence for the three data streams on the optimization, with atmospheric $CO_2$ data having the greatest average observation influence on the solution. A consequence of correcting the trend bias is that the model improvement with respect to other processes (photosynthesis, phenology) is hindered.

We attempted to overcome this by setting up a 2-step assimilation process where the trend correction is mostly achieved in the first step by tuning the regional parameters controlling the soil carbon pools. In doing so, the 2-step approach optimizes the constraint brought by *in situ* and satellite data (in the second step) in the joint assimilation process. Therefore, the 2-step results in enhanced model-data consistencies compared to a standard simultaneous assimilation (as observed in Figure 2 and Figure 3) with a caveat regarding atmospheric $CO_2$ data: the improved fit is mostly with the detrended atmospheric $CO_2$ data but not the raw data. We acknowledge the fact that this way of doing is not optimal and requires further investigation. But most importantly, it highlights the necessity to use atmospheric $CO_2$ data to constrain C cycle related parameters in order to improve the representation of soil carbon stocks in TBMs, which are pivotal to predicting NEE in regional to global assessments of the capacity of the terrestrial ecosystems to absorb or not atmospheric $CO_2$. Going beyond the steady state assumption following model spin-up has been discussed already (Carvalhais et al., (2010); MacBean et al., 2022), as it results in biased estimates of soil carbon reservoirs (Exbrayat et al., 2014). The exploitation of available datasets related to regional soil carbon stocks (as the International Soil Carbon Network, Nave et al. (2016), or the global soil respiration database, Jian et al. (2021)) by TBMs is not straightforward because of inconsistencies between the estimated quantities and the model state variables and underlying processes. Still, it is of primary

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022

Biogeosciences Discussions

importance for the science community to endeavor bridging the gap between state-of-the art estimates of soil carbon stocks and what TBMs simulate over the historical period.

### 4.3 Realism of the regional to global-scale C fluxes

The overarching objective of the study was more about assessing how to make the best of a synergistic exploitation of different data streams within a consistent assimilation framework rather than achieving an up-to-date re-analysis of the global carbon fluxes. However, the regional/global estimated NEE and GPP budgets are realistic and in agreement with independent estimates; but there are still important differences in the model predictions for the different assimilation experiments. We have not attempted to identify what was the most reliable optimized model, which would require the use of an ensemble of independent data (an effort beyond the scope of this paper). Interestingly, we noticed that the multi-data stream assimilations including $CO_2$ data (except the 2-step case) lead to a partition of the terrestrial C sink between the extra-tropics and the tropics in closer agreement to the atmospheric inversion than to the TBMs of the TRENDY experiment, a key feature that needs further investigations. In addition, we acknowledge that the coverage of the datasets used here is limited (no atmospheric $CO_2$ data nor satellite data after 2010, no *in situ* flux data beyond 2007) and that we did not assess the potential of other data that can bring relevant additional information on the dynamics of terrestrial carbon fluxes stocks, such as aboveground biomass (Thum et al., 2017) or Solar Induced-Fluorescence (Bacour et al., 2019) which have already been investigated with ORCHIDAS, and with an updated version of the ORCHIDEE model.

## 5 Conclusion

Terrestrial ecosystem modelers are anticipating the many novel types of observations that are being made available for model evaluation and assimilation. As a result, and in parallel to the growing complexity of TBMs incorporating new biogeo- physical processes related to the carbon and water cycles, new observation operators are being developed to be able to make use of this new wealth of data. We are thus entering an exciting new era, one in which we should both take advantage of a multitude of new datasets and capitalize on past gains in terrestrial model data assimilation understanding to achieve significant reduction in land surface model projection uncertainties.

We hope that the metrics explored in this study can benefit a broader set of data assimilation applications, supporting guidance for setting up the framework and for better use of the data to be assimilated.

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022

Biogeosciences
Discussions

## Code availability

The ORCHIDEE model code is open source (http://forge.ipsl.jussieu.fr/orchidee) and the associated documentation can be found at https://forge.ipsl.jussieu.fr/orchidee/wiki/Documentation. The ORCHIDAS data assimilation scheme (in Python) is available through a dedicated web site (https://orchidas.lsce.ipsl.fr/). Information about the LMDz model, source code and contact is provided at https://lmdz.lmd.jussieu.fr/le-projet-lmdz-en-bref-en.

## Data availability

This work used eddy covariance data acquired by the FLUXNET community (https://fluxnet.org/data/la-thuile-dataset/). The NDVI data are derived from the MODIS MOD09CMG collection 5 daily global reflectance products (‘https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/MOD09CMG). The surface atmospheric $CO_2$ concentration data uses measurements from The NOAA Earth System Laboratory (ESRL) archive (ftp://ftp.cmdl.noaa.gov/ccg/co2/), the CarboEurope IP project (http://ceatmosphere.lsce.ipsl.fr/database/index_database.html), and the World Data Centre for Greenhouse Gases of the World Meteorological Organization Global Atmospheric Watch Programme (http://gaw.kishou.go.jp).

## Author contributions

CB, NM, PP and FC conceived the research. CB developed the data assimilation system with contribution from FC (coupling with LMDz) and SL (parallelisation and post-processing). PP developed the offline transport (precomputed Jacobian matrix of LMDz) with contribution from SL. CB conducted the analysis, with contributions from NM and SL for spin-up ORCHIDEE simulations. PP, FC, and EK, provided the ancillary input fluxes for the global-scale simulations. EK and CB contributed to the development of the tangent linear version of the ORCHIDEE model. CB conceived and wrote the original draft with NM, PP, and FC. All co-authors reviewed the paper.

## Acknowledgements

This work has been supported by the CARBONES project, within the EU’s 7th Framework Program for Research and Development. The authors are very grateful to LSCE's IT staff for their support and for the computing resources, as well as to the ORCHIDEE Project Team for developing and maintaining the ORCHIDEE code.

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022



## References

Ahlström, A., Raupach, M. R., Schurgers, G., Smith, B., Arneth, A., Jung, M., Reichstein, M., Canadell, J. G., Friedlingstein, P., and Jain, A. K.: The dominant role of semi-arid ecosystems in the trend and variability of the land CO2 sink, 348, 895–899, 2015.

Arora, V. K., Katavouta, A., Williams, R. G., Jones, C. D., Brovkin, V., Friedlingstein, P., Schwinger, J., Bopp, L., Boucher, O., and Cadule, P.: Carbon–concentration and carbon–climate feedbacks in CMIP6 models and their comparison to CMIP5 models, 17, 4173–4222, 2020.

Bacour, C., Peylin, P., MacBean, N., Rayner, P. J., Delage, F., Chevallier, F., Weiss, M., Demarty, J., Santaren, D., and Baret, F.: Joint assimilation of eddy covariance flux measurements and FAPAR products over temperate forests within a process-oriented biosphere model, 120, 1839–1857, 2015.

Bacour, C., Maignan, F., Peylin, P., Macbean, N., Bastrikov, V., Joiner, J., Köhler, P., Guanter, L., and Frankenberg, C.: Differences between OCO-2 and GOME-2 SIF products from a model-data fusion perspective, 124, 3143–3157, 2019.

Bastrikov, V., MacBean, N., Bacour, C., Santaren, D., Kuppel, S., and Peylin, P.: Land surface model parameter optimisation using in situ flux data: comparison of gradient-based versus random search algorithms (a case study using ORCHIDEE v1. 9.5. 2), 11, 4739–4754, 2018.

Botta, A., Viovy, N., Ciais, P., Friedlingstein, P., and Monfray, P.: A global prognostic scheme of leaf onset using satellite data, 6, 709–725, 2000.

Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C.: A limited memory algorithm for bound constrained optimization, 16, 1190–1208, 1995.

Cardinali, C., Pezzulli, S., and Andersson, E.: Influence-matrix diagnostic of a data assimilation system, 130, 2767–2786, 2004.

Carvalhais, N., Reichstein, M., Ciais, P., Collatz, G. J., Mahecha, M. D., Montagnani, L., Papale, D., Rambal, S., and Seixas, J.: Identification of vegetation and soil carbon pools out of equilibrium in a process model via eddy covariance and biometric constraints, 16, 2813–2829, 2010.

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022




Cressot, C., Chevallier, F., Bousquet, P., Crevoisier, C., Dlugokencky, E. J., Fortems-Cheiney, A., Frankenberg, C., Parker, R., Pison, I., and Scheepmaker, R. A.: On the consistency between global and regional methane emissions inferred from SCIAMACHY, TANSO-FTS, IASI and surface measurements, 14, 577–592, 2014.

Crowell, S., Baker, D., Schuh, A., Basu, S., Jacobson, A. R., Chevallier, F., Liu, J., Deng, F., Feng, L., and McKain, K.: The 2015–2016 carbon cycle as seen from OCO-2 and the global in situ network, 19, 9797–9831, 2019.

Dee, D. P.: Bias and data assimilation, 131, 3323–3343, 2005.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., and Bauer, d P.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, 137, 553–597, 2011.

Desroziers, G., Berre, L., Chapnik, B., and Poli, P.: Diagnosis of observation, background and analysis-error statistics in observation space, 131, 3385–3396, 2005.

Dufresne, J.-L., Foujols, M.-A., Denvil, S., Caubel, A., Marti, O., Aumont, O., Balkanski, Y., Bekki, S., Bellenger, H., and Benshila, R.: Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5, 40, 2123–2165, 2013.

Exbrayat, J.-F., Pitman, A. J., and Abramowitz, G.: Response of microbial decomposition to spin-up explains CMIP5 soil carbon range until 2100, 7, 2683–2692, 2014.

Forkel, M., Carvalhais, N., Schaphoff, S., Migliavacca, M., Thurner, M., and Thonicke, K.: Identifying environmental controls on vegetation greenness phenology through model–data integration, 11, 7025–7050, 2014.

Fox, A. M., Hoar, T. J., Anderson, J. L., Arellano, A. F., Smith, W. K., Litvak, M. E., MacBean, N., Schimel, D. S., and Moore, D. J.: Evaluation of a data assimilation system for land surface models using CLM4. 5, 10, 2471–2494, 2018.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S., Aragão, L. E. O. C., Arneth, A., Arora, V., Bates, N. R., Becker, M., Benoit-Cattin, A., Bittig, H. C., Bopp, L., Bultan, S., Chandra, N., Chevallier, F., Chini, L. P., Evans, W., Florentie, L., Forster, P. M., Gasser, T., Gehlen, M., Gilfillan, D., Gkritzalis, T., Gregor, L., Gruber, N., Harris, I., Hartung, K., Haverd, V., Houghton, R. A., Ilyina, T., Jain, A. K., Joetzjer, E., Kadono, K., Kato, E., Kitidis, V., Korsbakken, J. I.,

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022




Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Liu, Z., Lombardozzi, D., Marland, G., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Niwa, Y., O'Brien, K., Ono, T., Palmer, P. I., Pierrot, D., Poulter, B., Resplandy, L., Robertson, E., Rödenbeck, C., Schwinger, J., Séférian, R., Skjelvan, I., Smith, A. J. P., Sutton, A. J., Tanhua, T., Tans, P. P., Tian, H., Tilbrook, B., van der Werf, G., Vuichard, N., Walker, A. P., Wanninkhof, R., Watson, A. J., Willis, D., Wiltshire, A. J., Yuan, W., Yue, X., and Zaehle, S.: Global Carbon Budget 2020, 12, 3269–3340, https://doi.org/10.5194/essd-12-3269-2020, 2020.

Giering, R., Kaminski, T., and Slawig, T.: Generating efficient derivative code with TAF: Adjoint and tangent linear Euler flow around an airfoil, 21, 1345–1355, 2005.

Groenendijk, M., Dolman, A. J., Van Der Molen, M. K., Leuning, R., Arneth, A., Delpierre, N., Gash, J. H. C., Lindroth, A., Richardson, A. D., and Verbeeck, H.: Assessing parameter variability in a photosynthesis model within and between plant functional types using global Fluxnet eddy covariance data, 151, 22–38, 2011.

Hourdin, F., Musat, I., Bony, S., Braconnot, P., Codron, F., Dufresne, J.-L., Fairhead, L., Filiberti, M.-A., Friedlingstein, P., and Grandpeix, J.-Y.: The LMDZ4 general circulation model: climate performance and sensitivity to parametrized physics with emphasis on tropical convection, 27, 787–813, 2006.

Jian, J., Vargas, R., Anderson-Teixeira, K., Stell, E., Herrmann, V., Horn, M., Kholod, N., Manzon, J., Marchesi, R., and Paredes, D.: A restructured and updated global soil respiration database (SRDB-V5), 13, 255–267, 2021.

Kaminski, T., Knorr, W., Rayner, P. J., and Heimann, M.: Assimilating atmospheric data into a terrestrial biosphere model: A case study of the seasonal cycle, 16, 14–1, 2002.

Kaminski, T., Knorr, W., Schürmann, G., Scholze, M., Rayner, P. J., Zaehle, S., Blessing, S., Dorigo, W., Gayler, V., and Giering, R.: The BETHY/JSBACH carbon cycle data assimilation system: Experiences and challenges, 118, 1414–1426, 2013.

Kato, T., Knorr, W., Scholze, M., Veenendaal, E., Kaminski, T., Kattge, J., and Gobron, N.: Simultaneous assimilation of satellite and eddy covariance data for improving terrestrial water and carbon simulations at a semi-arid woodland site in Botswana, 10, 789–802, 2013.

Keenan, T. F., Davidson, E. A., Munger, J. W., and Richardson, A. D.: Rate my data: quantifying the value of ecological data for the development of models of the terrestrial carbon cycle, 23, 273–286, 2013.

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022


Knorr, W. and Heimann, M.: Impact of drought stress and other factors on seasonal land biosphere CO2 exchange studied through an atmospheric tracer transport model, 47, 471–489, https://doi.org/10.1034/j.1600-0889.47.issue4.7.x, 1995.

Knorr, W. and Kattge, J.: Inversion of terrestrial ecosystem model parameter values against eddy covariance measurements by Monte Carlo sampling, 11, 1333–1351, 2005.

Knorr, W., Kaminski, T., Scholze, M., Gobron, N., Pinty, B., Giering, R., and Mathieu, P.-P.: Carbon cycle data assimilation with a generic phenology model, 115, 2010.

Koffi, E. N., Rayner, P. J., Scholze, M., and Beer, C.: Atmospheric constraints on gross primary productivity and net ecosystem productivity: Results from a carbon-cycle data assimilation system, 26, 2012.

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, 19, 2005.

Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F., and Richardson, A. D.: Constraining a global ecosystem model with multi-site eddy-covariance data, 9, 3757–3776, 2012.

Kuppel, S., Chevallier, F., and Peylin, P.: Quantifying the model structural error in carbon cycle data assimilation systems, Geosci. Model Dev., 6, 45–55, doi: 10.5194, gmd-6-45-2013, 2013.

Kuppel, S., Peylin, P., Maignan, F., Chevallier, F., Kiely, G., Montagnani, L., and Cescatti, A.: Model–data fusion across ecosystems: from multisite optimizations to global simulations, 7, 2581–2597, 2014.

Luo, Y. Q., Randerson, J. T., Abramowitz, G., Bacour, C., Blyth, E., Carvalhais, N., Ciais, P., Dalmonech, D., Fisher, J. B., Fisher, R., Friedlingstein, P., Hibbard, K., Hoffman, F., Huntzinger, D., Jones, C. D., Koven, C., Lawrence, D., Li, D. J., Mahecha, M., Niu, S. L., Norby, R., Piao, S. L., Qi, X., Peylin, P., Prentice, I. C., Riley, W., Reichstein, M., Schwalm, C., Wang, Y. P., Xia, J. Y., Zaehle, S., and Zhou, X. H.: A framework for benchmarking land models, 9, 3857–3874, https://doi.org/10.5194/bg-9-3857-2012, 2012.

MacBean, N., Maignan, F., Peylin, P., Bacour, C., Bréon, F.-M., and Ciais, P.: Using satellite data to improve the leaf phenology of a global terrestrial biosphere model, 12, 7185–7208, 2015.

MacBean, N., Peylin, P., Chevallier, F., Scholze, M., and Schürmann, G.: Consistent assimilation of multiple data streams in a carbon cycle data assimilation system, 9, 3569–3588, 2016.

MacBean, N., Bacour, C., Raoult, N., Bastrikov, V., Koffi, E. N., Kuppel, S., Maignan, F., Ottlé, C., Peaucelle, M., Santaren, D., and Peylin, P.: Quantifying and Reducing Uncertainty in Global Carbon Cycle Predictions: Lessons and Perspectives From 15 Years of Data Assimilation Studies with the ORCHIDEE Terrestrial Biosphere Model, Global Biogeochemical Cycles, https://doi.org/10.1029/2021GB007177

Migliavacca, M., Meroni, M., Busetto, L., Colombo, R., Zenone, T., Matteucci, G., Manca, G., and Seufert, G.: Modeling gross primary production of agro-forestry ecosystems by assimilation of satellite-derived information in a process-based model, 9, 922–942, 2009.

Moore, D. J., Hu, J., Sacks, W. J., Schimel, D. S., and Monson, R. K.: Estimating transpiration and the sensitivity of carbon uptake to water availability in a subalpine forest using a simple ecosystem process model informed by measured net CO2 and H2O fluxes, 148, 1467–1477, 2008.

Nave, L., Johnson, K., van Ingen, C., Agarwal, D., Humphrey, M., and Beekwilder, N.: International Soil Carbon Network (ISCN) Database v3-1, International Soil Carbon Network (ISCN), 2016.

Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., and Vesala, T.: Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation, 3, 571–583, 2006.

Parton, W. J., Schimel, D. S., Cole, C. V., and Ojima, D. S.: Analysis of factors controlling soil organic matter levels in Great Plains grasslands, 51, 1173–1179, 1987.

Peiro, H., Crowell, S., Schuh, A., Baker, D. F., O'Dell, C., Jacobson, A. R., Chevallier, F., Liu, J., Eldering, A., and Crisp, D.: Four years of global carbon cycle observed from the Orbiting Carbon Observatory 2 (OCO-2) version 9 and in situ data and comparison to OCO-2 version 7, 22, 1097–1130, 2022.

Peylin, P., Law, R. M., Gurney, K. R., Chevallier, F., Jacobson, A. R., Maki, T., Niwa, Y., Patra, P. K., Peters, W., and Rayner, P. J.: Global atmospheric carbon budget: results from an ensemble of atmospheric CO 2 inversions, 10, 6699–6720, 2013.

Peylin, P., Bacour, C., MacBean, N., Leonard, S., Rayner, P., Kuppel, S., Koffi, E., Kane, A., Maignan, F., and Chevallier, F.: A new stepwise carbon cycle data assimilation system using multiple data streams to constrain the simulated land surface carbon cycle, 9, 2016.

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022



Quaife, T., Lewis, P., De Kauwe, M., Williams, M., Law, B. E., Disney, M., and Bowyer, P.: Assimilating canopy reflectance data into an ecosystem model with an Ensemble Kalman Filter, 112, 1347–1364, 2008.

Randerson, J. T., van der Werf, G. R., Giglio, L., Collatz, G. J., and Kasibhatla, P. S.: Global Fire Emissions Database, Version 3 (GFEDv3. 1), Data set, Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA, 2013.

Raoult, N. M., Jupp, T. E., Cox, P. M., and Luke, C. M.: Land surface parameter optimisation through data assimilation: the ad-JULES system, Geosci. Model Dev. Discuss., doi: 10.5194, 2016.

Raupach, M. R., Rayner, P. J., Barrett, D. J., DeFries, R. S., Heimann, M., Ojima, D. S., Quegan, S., and Schmullius, C. C.: Model–data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications, 11, 378–397, 2005.

Rayner, P. J., Scholze, M., Knorr, W., Kaminski, T., Giering, R., and Widmann, H.: Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS), 19, 2005.

Ricciuto, D. M., King, A. W., Dragoni, D., and Post, W. M.: Parameter and prediction uncertainty in an optimized terrestrial carbon cycle model: Effects of constraining variables and data record length, 116, 2011.

Richardson, A. D., Williams, M., Hollinger, D. Y., Moore, D. J., Dail, D. B., Davidson, E. A., Scott, N. A., Evans, R. S., Hughes, H., and Lee, J. T.: Estimating parameters of a forest ecosystem C model with measurements of stocks and fluxes as joint constraints, 164, 25–40, 2010.

Sacks, W. J., Schimel, D. S., and Monson, R. K.: Coupling between carbon cycling and climate in a high-elevation, subalpine forest: a model-data fusion analysis, 151, 54–68, 2007.

Santaren, D., Peylin, P., Viovy, N., and Ciais, P.: Optimizing a process-based ecosystem model with eddy-covariance flux measurements: A pine forest in southern France, 21, 2007.

Santaren, D., Peylin, P., Bacour, C., Ciais, P., and Longdoz, B.: Ecosystem model optimization using in situ flux observations: benefit of Monte Carlo versus variational schemes and analyses of the year-to-year model performances, 11, 7137–7158, 2014.

Schürmann, G. J., Kaminski, T., Köstler, C., Carvalhais, N., Vo\s sbeck, M., Kattge, J., Giering, R., Rödenbeck, C., Heimann, M., and Zaehle, S.: Constraining a land-surface model with multiple

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022

Biogeosciences Discussions
Open Access

observations by application of the MPI-Carbon Cycle Data Assimilation System V1. 0, Geosci. Model Dev., 9, 2999–3026, gmd-9-2999-2016, 2016.

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S. D., Murray-Tortarolo, G., Ahlström, A., Doney, S. C., Graven, H., Heinze, C., and Huntingford, C.: Recent trends and drivers of regional sources and sinks of carbon dioxide, 12, 653–679, 2015.

Stöckli, R., Rutishauser, T., Dragoni, D., O'keefe, J., Thornton, P. E., Jolly, M., Lu, L., and Denning, A. S.: Remote sensing data assimilation for a prognostic phenology model, 113, 2008.

Thum, T., MacBean, N., Peylin, P., Bacour, C., Santaren, D., Longdoz, B., Loustau, D., and Ciais, P.: The potential benefit of using forest biomass data in addition to carbon and water flux measurements to constrain ecosystem model parameters: case studies at two temperate forest sites, 234, 48–65, 2017.

Vermote, E., Justice, C. O., and Bréon, F.-M.: Towards a generalized approach for correction of the BRDF effect in MODIS directional reflectances, 47, 898–908, 2008.

Wang, Y.-P., Leuning, R., Cleugh, H. A., and Coppin, P. A.: Parameter estimation in surface exchange models using nonlinear inversion: how many parameters can we estimate and which measurements are most useful?, 7, 495–510, 2001.

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Kasibhatla, P. S., and Arellano Jr, A. F.: Interannual variability in global biomass burning emissions from 1997 to 2004, 6, 3423–3441, 2006.

Williams, M., Richardson, A. D., Reichstein, M., Stoy, P. C., Peylin, P., Verbeeck, H., Carvalhais, N., Jung, M., Hollinger, D. Y., and Kattge, J.: Improving land surface models with FLUXNET data, 6, 1341–1359, 2009.

Wutzler, T. and Carvalhais, N.: Balancing multiple constraints in model-data integration: Weights and the parameter block approach, 119, 2112–2129, https://doi.org/10.1002/2014JG002650, 2014.

Zhu, C., Byrd, R. H., Lu, P., and Nocedal, J.: Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization, 23, 550–560, 1997.

Zobitz, J. M., Moore, D. J., Quaife, T., Braswell, B. H., Bergeson, A., Anthony, J. A., and Monson, R. K.: Joint data assimilation of satellite reflectance and net ecosystem exchange data constrains ecosystem carbon fluxes at a high-elevation subalpine forest, 195, 73–88, 2014.

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022


Biogeosciences
Discussions
Open Access

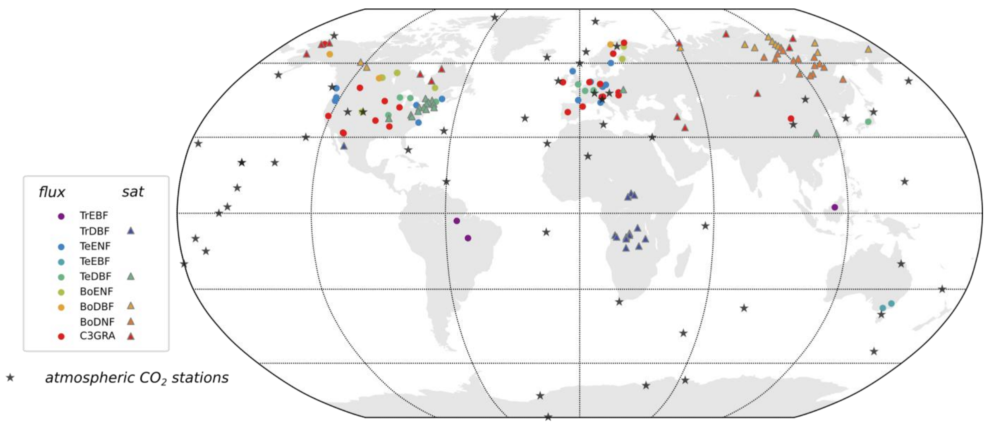


**Figure 1: Location of the flux tower sites (circles), satellite pixels (triangles), and atmospheric $CO_2$ stations (black stars) used in this study.**

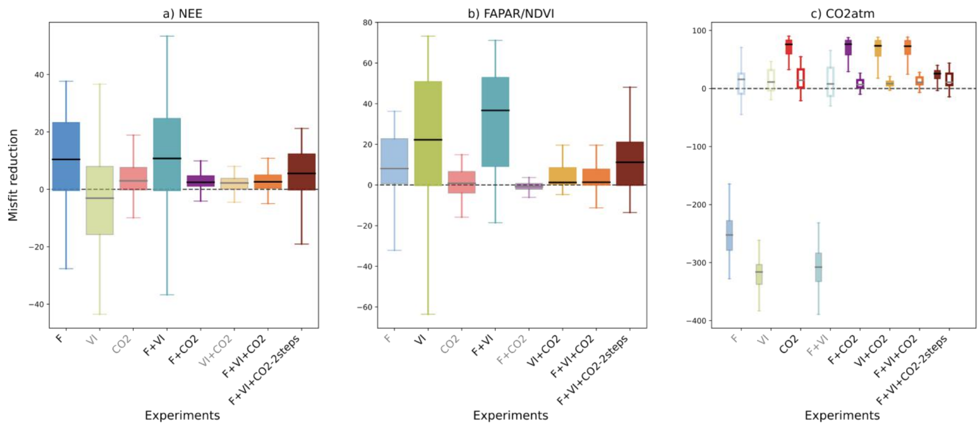


**Figure 2: For all data streams, boxplots of the reduction of the model-data mismatch following the different assimilation experiments. For a given data stream, the assimilation experiments in which it is involved are labeled in black (x-axis) and the boxplot colors are dark colored; and in gray / light colors otherwise (back-compatibility check). For the atmospheric $CO_2$ concentration data at stations, the misfit reduction is calculated both for the raw (not detrended) data (left solid boxplot of each assimilation experiment, with colored boxplots) and the detrended data (right white boxplot of each assimilation experiment).**

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022



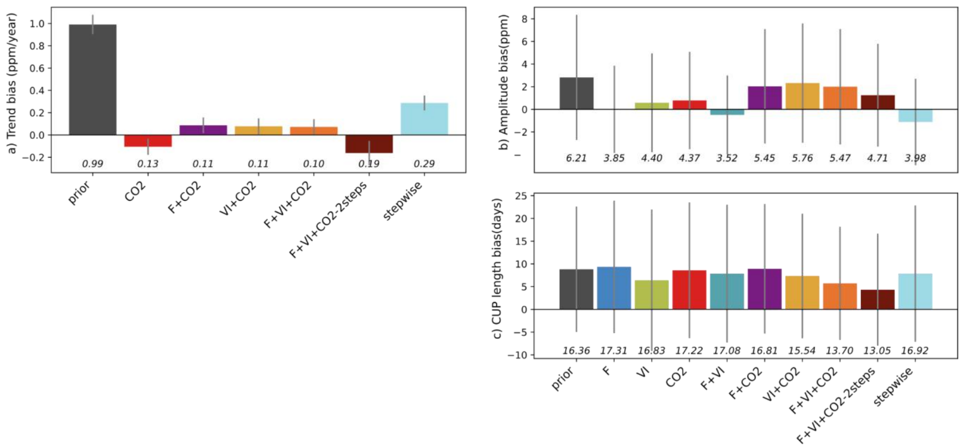


**Figure 3: Residual biases of the atmospheric $CO_2$ time series between those measured at stations and the simulations (prior and optimized for each assimilation experiment), in terms of trend, magnitude of the seasonal cycle and length of the carbon uptake (CUP). The study results are compared to those obtained using a sequential approach (Peylin et al., 2016). The bars show for each quantity the mean bias relative to the measurements over the period 2000-2009. The standard deviations of the differences between observations and simulations over all stations are shown as the gray vertical lines, and the RMSD are provided below in italic.**

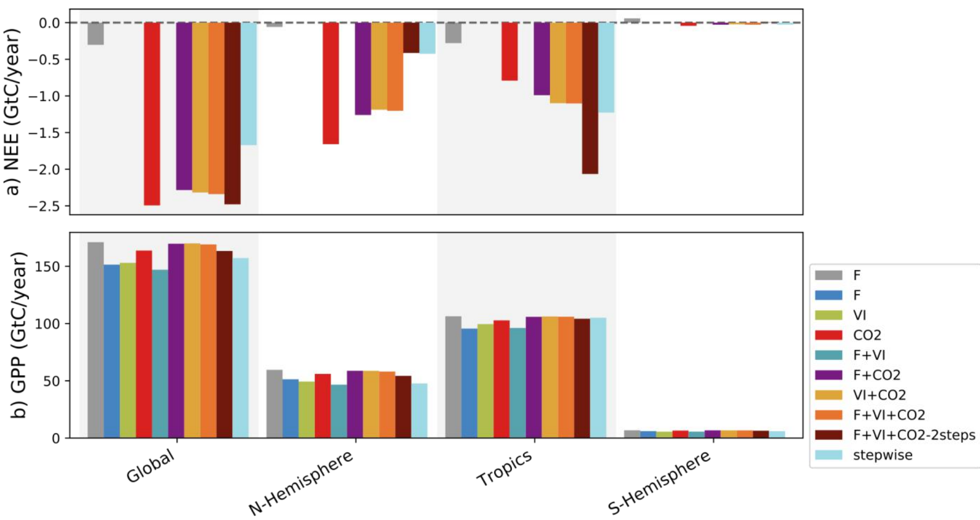


**Figure 4: Global and regional C budget for NEE and GPP, and for the northern hemisphere (30°N-90°N), tropics (30°N-30°S) and southern hemisphere (30°S-90°S), regions, for the prior model and the model calibrated for the several assimilation experiments. For NEE, only the experiments involving atmospheric $CO_2$ data are shown. The period considered is 2000-2009.**

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022




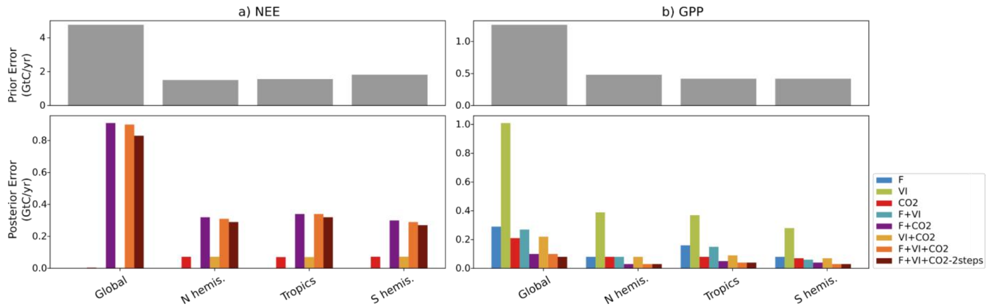


**Figure 5: For NEE (left) and GPP (right) prior errors (top), and posterior errors obtained for each assimilation experiment (bottom), over the regions considered. For NEE, only the experiments involving atmospheric $CO_2$ data are shown.**

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022




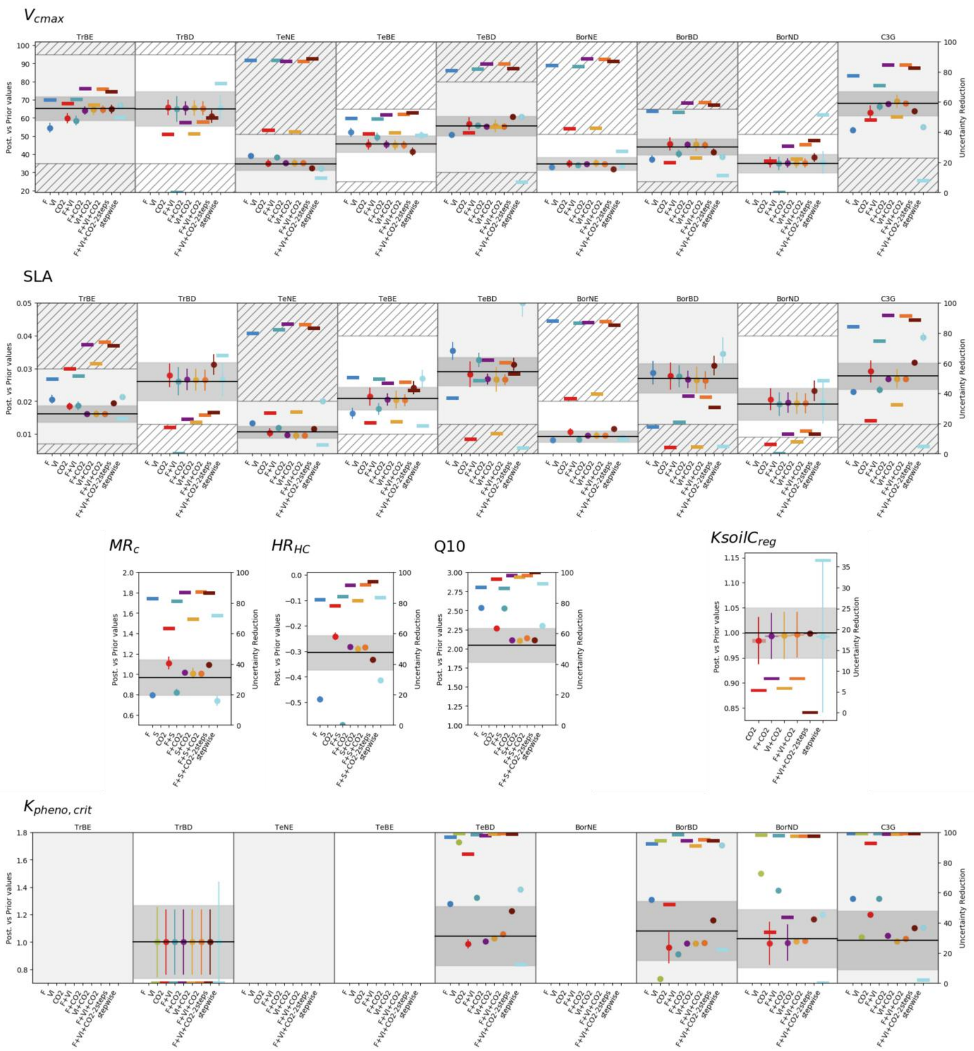


**Figure 6: Prior and posterior parameter values and uncertainties for a set of optimized parameters (eight PFT-dependent parameters and four non-PFT dependent). The prior value is shown as the horizontal black line and the prior uncertainty (standard deviation) as the gray area encompassing it along the x-axis. For the PFT-dependent parameters, each box corresponds to a given PFT; empty boxes indicate that this parameter was not constrained for the corresponding PFTs. The white zone (non-dashed area) corresponds to the allowed range of variation. The optimized values are provided for each assimilation experiment (the eight ones considered in this study and the one from Peylin et al. (2016) – "stepwise"); the corresponding posterior errors are displayed as the vertical bars. Note that the prior values presented here are those used in this study, and not those of the stepwise (which are higher/lower for the photosynthesis and respiration / phenological parameters). For each assimilation experiment is also provided the uncertainty reduction (right y-axis) as the thick opaque horizontal bars. For *KsoilC_reg*, the posterior values displayed here correspond to**

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022


Biogeosciences
Discussions
Open Access

**the mean over the ecoregions (without Antarctica) considered; the semi-transparent horizontal bars on either side of the posterior values correspond to the standard deviation of the estimates.**

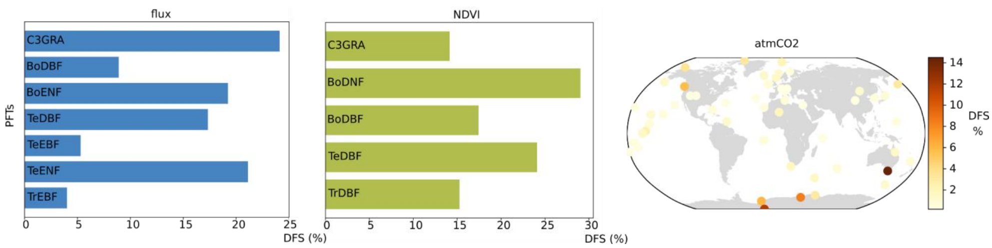


**Figure 7: Relative DFS for the F+VI+CO2 assimilation experiment. For Flux and Satellite data: relative DFS per PFT; for atmospheric $CO_2$ data: relative relative DFS (contribution) of the different stations to the fit.**

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022



| Name | Description | Data stream |
|---|---|---|
| *Photosynthesis* | | |
| **$V_{cmax}$** | maximum carboxylation rate (µmol.m$^{-2}$.s$^{-1}$) | F, CO2 |
| **$G_{s,slope}$** | Ball-Berry slope | F, CO2 |
| **$T_{opt}$** | optimal photosynthesis temperature (°C) | F, CO2 |
| **SLA** | specific leaf area (m$^2$.g$^{-1}$) | F, CO2 |
| *Soil water availability* | | |
| **$H_{um,cste}$** | root profile (m$^{-1}$) | F, CO2 |
| *Phenology* | | |
| **$LAI_{MAX}$** | maximum LAI value | F, CO2 |
| **$K_{pheno,crit}$** | multiplicative parameter of the threshold that determines the start of the growing season | F, VI, CO2 |
| **$T_{senes}$** | temperature threshold for senescence (°C) | F, VI, CO2 |
| **$L_{age,crit}$** | average critical age of leaves (days) | F, VI, CO2 |
| **$K_{LAI,happy}$** | LAI threshold to stop using carbohydrate reserves | F, VI, CO2 |
| *Respiration* | | |
| **Q10** | temperature dependency of heterotrophic respiration | F, CO2 |
| **$HR_{H,c}$** | Offset of the function for moisture control factor of heterotrophic respiration | F, CO2 |
| **$MR_c$** | Offset of the affine relationship between temperature and maintenance respiration | F, CO2 |
| **$K_{soilC,site}$** | Multiplicative factor of initial slow and passive carbon pools | F |
| **$K_{soilC,reg}$** | Multiplicative factor of initial slow and passive carbon pools | CO2 |

**Table 1: List of the ORCHIDEE parameters to be optimized and data streams that constrain them (F for *in situ* flux measurements, VI for normalized satellite NDVI data, CO2 for atmospheric $CO_2$ concentration data).**

| experiment name | flux data | NDVI data | atmospheric $CO_2$ concentrations | number of optimized parameters | number of observations |
|---|---|---|---|---|---|
| **F** | x | | | 133 | 150792 |
| **VI** | | x | | 19 | 149916 |
| **CO2** | | | x | 114 | 6360 |
| **F+VI** | x | x | | 152 | 300708 |
| **F+CO2** | x | | x | 182 | 157152 |
| **VI+CO2** | | x | x | 114 | 156276 |
| **F+VI+CO2**<br>**F+VI+CO2-2steps** | x | x | x | 182 | 307068 |

**Table 2: Characteristics of the various assimilation experiments (flux data – F, satellite NDVI vegetation index – VI, and atmospheric CO2 concentration – CO2).**

https://doi.org/10.5194/bg-2022-109
Preprint. Discussion started: 31 May 2022




| | NEE | LE | VI | CO2 |
|---|---|---|---|---|
| $\mathbf{R}$ | 1.75 | 1.75 | 0.33 | 1.22 |
| $E[\mathbf{d}^{\mathbf{o}}_{\mathbf{a}}.\mathbf{d}^{\mathbf{o}}_{\mathbf{b}}{}^{t}]$ | 1.49 | 1.49 | 0.21 | 1.16 |
| $ratio^{R}$ | 1.17 | 1.17 | 1.55 | 1.05 |
| | | | | |
| $\mathbf{H_o}.\mathbf{B}.\mathbf{H_o}{}^{t}$ | 1.45 | 8.30 | 0.2 | 15.17 |
| $E[\mathbf{d}^{\mathbf{a}}_{\mathbf{b}}.\mathbf{d}^{\mathbf{o}}_{\mathbf{b}}{}^{t}]$ | 0.92 | 5.45 | 0.24 | 6.29 |
| $ratio^{B}$ | 1.59 | 1.52 | 0.83 | 2.41 |
| | | | | |
| $\mathbf{H_o}.\mathbf{B}.\mathbf{H_o}{}^{t}+\mathbf{R}$ | 2.28 | 23.63 | 0.38 | 15.22 |
| $E[\mathbf{d}^{\mathbf{o}}_{\mathbf{b}}.\mathbf{d}^{\mathbf{o}}_{\mathbf{b}}{}^{t}]$ | 1.75 | 22.11 | 0.31 | 6.39 |
| $ratio^{BR}$ | 1.17 | 1.07 | 1.23 | 2.38 |
| | | | | |
| $\mathbf{H_o}.\mathbf{A}.\mathbf{H_o}{}^{t}$ | 0.25 | 1.82 | 0.07 | 3.26 |
| $E[\mathbf{d}^{\mathbf{a}}_{\mathbf{b}}.\mathbf{d}^{\mathbf{o}}_{\mathbf{a}}{}^{t}]$ | -0.45 | -5.12 | -0.15 | -2.13 |
| $ratio^{A}$ | -0.56 | -0.36 | -0.43 | -1.53 |

**Table 3: Consistency diagnostics of the error covariance matrices for the F (using NEE and LE data), VI, and CO2, assimilation experiments. The ratios are calculated with the mathematical expectation term as the denominator.**

| | OI | | Relative DFS | |
|---|---|---|---|---|
| | **1-step** | **2-step** | **1-step** | **2-step** |
| **flux** | 0.000586 | 0.000577 | 74.65 | 76.9 |
| **NDVI** | 0.000048 | 0.000048 | 11.12 | 11.68 |
| **CO2** | 0.002654 | 0.002035 | 14.23 | 11.42 |

**Table 4: Observation influence and relative DFS statistics of each data stream for the joint assimilation experiments F+VI+CO2 and F+VI+CO2-2steps.**